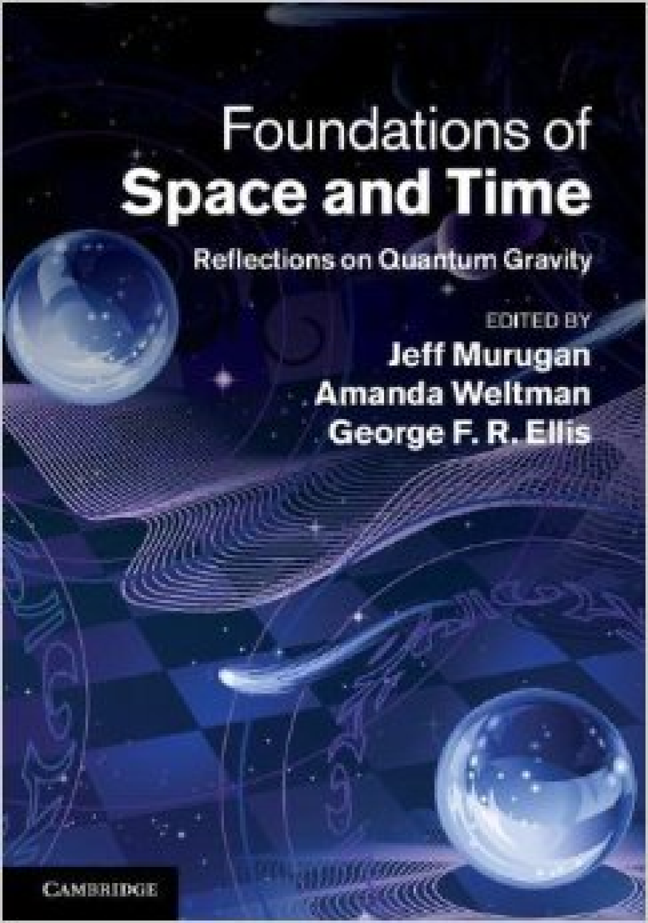

Foundations of
Space and Time
Reflections on Quantum Gravity
EDITED BY
Jeff Murugan
Amanda Weltman
George F. R. Ellis
CAMBRIDGE

FOUNDATIONS OF SPACE AND TIME

Reflections on Quantum Gravity

After almost a century, the field of quantum gravity remains as difficult and inspiring as ever. Today, it finds itself a field divided, with two major contenders dominating: string theory, the leading exemplification of the covariant quantization program; and loop quantum gravity, the canonical scheme based on Dirac's constrained Hamiltonian quantization. However, there are now a number of other innovative schemes providing promising new avenues.

Encapsulating the latest debates on this topic, this book details the different approaches to understanding the very nature of space and time. It brings together leading researchers in each of these approaches to quantum gravity to explore these competing possibilities in an open way. Its comprehensive coverage explores all the current approaches to solving the problem of quantum gravity, addressing the strengths and weaknesses of each approach, to give researchers and graduate students an up-to-date view of the field.

JEFF MURUGAN is a Senior Lecturer in the Department of Mathematics and Applied Mathematics and a member of the Astrophysics, Cosmology & Gravity Center, University of Cape Town. He is interested in all aspects of gravity and is currently working on string theory and connections between gauge theories and gravity.

AMANDA WELTMAN is a Senior Lecturer in the Department of Mathematics and Applied Mathematics and a member of the Astrophysics, Cosmology & Gravity Center, University of Cape Town. She works in the exciting bridging areas of string cosmology, studying physical ways to test string theory within the context of cosmology.

GEORGE F. R. ELLIS is Emeritus Professor of Applied Mathematics and Honorary Research Associate in the Mathematics Department, University of Cape Town. He works on general relativity theory, cosmology, complex systems, and the way physics underlies the functioning of the human brain.

FOUNDATIONS OF SPACE AND TIME

Reflections on Quantum Gravity

Edited by

JEFF MURUGAN, AMANDA WELTMAN &
GEORGE F. R. ELLIS

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521114400

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Foundations of space and time : reflections on quantum gravity / [edited by] Jeff Murugan,
Amanda Weltman & George F. R. Ellis.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-11440-0 (hardback)
1. Space and time. 2. Quantum gravity. I. Murugan, Jeff. II. Weltman, Amanda.
III. Ellis, George F. R. (George Francis Rayner) IV. Title.
QC173.59.S65F68 2011
531′.14–dc22 2011000387

ISBN 978-0-521-11440-0 Hardback

Contents

Contributors

J. Ambjørn
The Niels Bohr Institute, Copenhagen University,
Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark
and
Institute for Theoretical Physics, Utrecht University,
Leuvenlaan 4, NL-3584 CE Utrecht, The Netherlands

Massimo Bianchi
Dipartimento di Fisica and Sezione I.N.F.N.,
Università di Roma "Tor Vergata",
Via della Ricerca Scientifica, 00133 Roma, Italy

Martin Bojowald
Institute for Gravitation and the Cosmos,
Penn State University,
State College, PA 16801, USA

Cliff Burgess
Department of Physics & Astronomy, McMaster University,
1280 Main St. W, Hamilton, Ontario, Canada, L8S 4M1
and
Perimeter Institute for Theoretical Physics,
31 Caroline St. N,
Waterloo, Ontario, Canada, N2L 2Y5

Steven Carlip
Department of Physics,
University of California,
Davis, CA 95616, USA

Robert de Mello Koch
National Institute for Theoretical Physics,
Department of Physics and Centre for Theoretical Physics,
University of the Witwatersrand, Wits, 2050, South Africa

George F. R. Ellis
Astrophysics, Cosmology & Gravity Center,
University of Cape Town, Private Bag,
Rondebosch, 7700, South Africa

Joe Henson
Perimeter Institute,
31 Caroline Street North,
Waterloo, Ontario, Canada, N2L 2Y5

J. Jurkiewicz
Jagiellonian University,
Krakow Institute of Physics
Reymonta 4
Krakow 30-059, Poland

Axel Kleinschmidt
Physique Théorique et Mathématique &
International Solvay Institutes, Université Libre de Bruxelles,
Boulevard du Triomphe, ULB-CP231,
BE-1050 Bruxelles, Belgium

Renate Loll
Institute for Theoretical Physics, Utrecht University,
Leuvenlaan 4, NL-3584 CE Utrecht, The Netherlands

Jeff Murugan
Astrophysics, Cosmology & Gravity Center,
University of Cape Town, Private Bag,
Rondebosch, 7700, South Africa

Hermann Nicolai
Max-Planck-Institut für Gravitationsphysik,
Albert-Einstein-Institut,
Am Mühlenberg 1, DE-14476, Golm, Germany

Daniele Oriti
Max-Planck-Institut für Gravitationsphysik,
Albert-Einstein-Institut,
Am Mühlenberg 1, DE-14476 Golm, Germany

Thanu Padmanabhan
IUCAA, Pune University Campus,
Ganeshkhind, Pune 411007, India

Roger Penrose
The Mathematical Institute,
24–29 Saint Giles, Oxford OX1 3LB, UK

Dimitri Polyakov
National Institute for Theoretical Physics,
Department of Physics and Centre for Theoretical Physics,
University of the Witwatersrand,
Wits, 2050, South Africa

Hanno Sahlmann
Aria Pacific Center for Theoretical Physics
Hogil Kim Memorial Bldg. POSTECH
San 31, Hyoga-dong, Nam-gu
Pohang 790-784, Republic of Korea

Rafael Sorkin
Perimeter Institute,
31 Caroline Street N., Waterloo, Ontario, Canada, N2L 2Y5
and
Department of Physics,
Syracuse University, Syracuse, NY 13244-1130, USA

Kellog Stelle
Imperial College of Science, Technology and Medicine,
London Physics Department
South Kensington Campus
London
SW7 2AZ

Y. Watabiki
Tokyo Institute of Technology,
Dept. of Physics, High Energy Theory Group,
2-12-1 Oh-okayama, Meguro-ku, Tokyo 152-8551, Japan

Amanda Weltman
Astrophysics, Cosmology & Gravity Center,
University of Cape Town, Private Bag,
Rondebosch, 7700, South Africa

W. Westra
Department of Physics, University of Iceland,
Dunhaga 3, 107 Reykjavik, Iceland

S. Zohren
Mathematical Institute, Leiden University,
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

1

The problem with quantum gravity

JEFF MURUGAN, AMANDA WELTMAN & GEORGE F.R. ELLIS

"The effort to understand the Universe is one of the very few things that lifts human life a little above the level of farce, and gives it some of the grace of tragedy."

Steven Weinberg, *The First Three Minutes*, 1997

After almost a century, the field of quantum gravity remains as difficult, frustrating, inspiring, and alluring as ever. Built on answering just one question – *How can quantum mechanics be merged with gravity?* – it has developed into the modern muse of theoretical physics.

Things were not always this way. Indeed, inspired by the monumental victory against the laws of Nature that was quantum electrodynamics (QED), the 1950s saw the frontiers of quantum physics push to the new and unchartered territory of gravity with a remarkable sense of optimism. After all, if nothing else, gravity is orders of magnitude weaker than the electromagnetic interaction; surely it would succumb more easily. Nature, it would seem, is not without a sense of irony. For an appreciation of how this optimism eroded over the next 30 years, there is perhaps no better account than Feynman's *Lectures on Gravitation*. Contemporary with his epic *Feynman Lectures on Physics*, these lectures document Feynman's program of quantizing gravity "like a field theorist." In it he sets out to reverse-engineer a theory of gravity starting from the purely phenomenological observations that gravity is a long-range, static interaction that couples to the energy content of matter with universal attraction. Taken together, these facts hint toward a field theory built from a massless, spin-2 graviton propagating on a flat, Minkowski background, i.e., $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. The question of quantizing gravity then distills down to how to formulate a consistent quantum theory of this graviton. The consequences

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

of this, quite simplistic, viewpoint are profound. For example, in the quantum theory, a massless spin-2 graviton has two helicity states and hence so too must the associated classical gravitational field have two dynamical degrees of freedom. For this counting to match, one is *forced* to incorporate a redundancy to encode that many possible classical field configurations could correspond to a single physical state, i.e., *a gauge symmetry*. Ultimately, it is this gauge symmetry that can be interpreted as the principle of equivalence in the low-energy, classical limit.

By all accounts, Feynman's foray into quantum gravity culminated in the early 1960s with a covariant quantization of the gravitational field to one-loop order. This begs the question then of why, 50 years on and with the general principles laid down, this pragmatic program of quantizing gravity has not reached completion? There are really two main problems with this quantization scheme. The first is as contemporary as they get but is really an age-old issue that goes all the way back to Einstein himself: *the cosmological constant*. Classically, there are no theoretical constraints on Λ and it can, without much ado, be set to zero. In a quantum field theory, however, every field has an infinite number of modes, each of which possess a "zero-point" energy. Consequently, one expects that the vacuum energy of the field is infinite. In flat space this problem is easily overcome by redefining the (arbitrary) zero-point of the energy scale. Gravity, on the other hand, couples to the *energy content* of a system so that when the gravitational interaction is turned on, the vacuum fluctuations of any quantized field generate actual physical effects. Moreover, even if the modes are cut off at some momentum scale, the vacuum energy density generated by the remaining modes can still be quite large, in stark contrast to all observations (about 123 orders of magnitude so, in fact).

The second problem is the equally thorny question of the *renormalizability* of the quantum theory. Although more technical, it can nevertheless be summarized very roughly as follows. Every loop in a covariant Feynman diagram expansion contributes

$$I_{\text{loop}} \sim \int d^D p \, p^{4J-8}$$

in D spacetime dimensions when the interaction is mediated by a spin-J particle. This contribution is finite when $4J-8+D<0$ and infinite when $4J-8+D>0$. In the marginal case where $4J-8+D=0$, loop contributions diverge but only logarithmically and can always be absorbed into a redefinition of various couplings in the theory. This is the case in a renormalizable theory. Gravity in four dimensions is mediated by a spin-2 boson, the graviton, and consequently receives infinite contributions at each loop order. In this case, an infinite number of parameters are required to absorb all of the divergences and the theory is non-renormalizable. An equivalent way of phrasing this is in terms of the coupling constant. In units of $\hbar=c=1$, any theory whose coupling constant has a positive mass-dimension is finite. If the coupling constant is either dimensionless or has negative mass-dimension then the theory is renormalizable or non-renormalizable respectively. In general relativity, the coupling constant, G_N, has mass-dimension -2 and, again, the theory is non-renormalizable.[1] Nevertheless, this perturbative covariant approach historically illuminated the way forward.

Essentially, the more symmetric a theory is, the more tightly constrained are the counter-terms generated by the renormalization process and consequently, the more convergent it will be. Apparently then, one way to improve the ultraviolet behavior of a theory is to build more symmetry into it. This of course is the line of reasoning that led, in the 1970s, to the idea of *supergravity*, a theory of local (or gauged) supersymmetry that mixes bosonic and fermionic fields in a way that necessarily incorporates general covariance and hence gravity. The ultraviolet behavior of the quantum theory is under better control essentially because divergent bosonic (fermionic) loop contributions are cancelled by the associated contribution coming from the fermionic (bosonic) super-partner. For a time these supergravity theories (and the $\mathcal{N} = 8$ theory in four dimensions in particular) provided an enormous source of comfort for a community still reeling from prolonged battles against the infinities of quantum gravity.[2] However, it was soon realized that, even this much enlarged symmetry could not guarantee finiteness at all orders in the loop expansion and the supergravity machine lost a lot of its momentum.[3] Fortunately, another juggernaut loomed on the horizon.

Touted variously as the most promising candidate for a theory of quantum gravity, the "only game in town," or even the "theory of everything," string theory is a quantum theory of one-dimensional objects whose size is Planckian and whose different oscillation modes constitute the different members of the particle zoo. In particular, the first excited mode of a quantum closed string is a massless, spin-2 state that is identified with a graviton. String theory then appears to be a mathematically consistent (anomaly-free) quantum theory of gravity but, and perhaps more

[1] In contemporary terms, this non-renormalizability can be understood as a result of the fact that general relativity is an *effective field theory*, encoding low-energy gravitational dynamics (as summed up beautifully in Chapters 2 and 3). At small scales and high energies, this effective treatment breaks down and can manifest in a number of rather interesting phenomena. One such phenomenon, a change in the number of dimensions of space, can be found in Carlip's study of the small-scale structure of spacetime in this volume.

[2] So much so, in fact, that in his inaugural lecture for the Lucasian chair, Stephen Hawking declared that $\mathcal{N} = 8$ supergravity might just be the final theory signaling the "end of theoretical physics"!

[3] This momentum has resurfaced with a vengeance in the past few months (of editing this book). Following from an astounding observation of Witten on the relation between perturbative string theory and perturbative gauge theory formulated in twistor variables, a remarkable new insight appeared about the structure of gluon scattering amplitudes. When combined with the Kawai–Lewellen–Tye relations, this provided just the ammunition needed to resume the assault on the finiteness problem of $\mathcal{N} = 8$ supergravity. Indeed, initial reports from the front seem quite positive (see the discussions by Stelle and Nicolai in this volume).

importantly, it also *necessarily* contains quantum versions of the remaining fundamental interactions. Here, for the first time, was a theory where one was *forced* to consider *all* the fundamental forces of nature at once. However, famously, even after 30 years of painstaking work, string theory remains incomplete.

The problems with string theory are manifold. Historically, the first one to emerge was its dimensionality. In string theory "dimension" is no longer a fixed concept. It is instead a property of particular solutions of the theory. For example, any anomaly- and ghost-free solution of the superstring equations of motion possessing $\mathcal{N}=1$ supersymmetry on the worldsheet must have a spacetime dimension[4] $D=10$. While this problem can be circumvented by the old idea of Kaluza–Klein compactification it leads directly to the more thorny question of the uniqueness of solutions of the theory. Each compactification leads to a different vacuum state of string theory and since, if it is correct, at least one such state should describe our Universe in its entirety, the potentially enormous number ($\sim 10^{500}$ at last count) of consistent solutions, with no perturbative mechanism to select among them,[5] leads some critics to question the predictive power of the theory. Even more worrying is the fact that, while the theory is perturbatively finite in the sense discussed above (i.e., order by order), the perturbation series does not appear to converge. The veracity of the claims of finiteness of the theory is consequently unclear. By the time the 1990s rolled around the field found itself, somewhat understandably, in a state of malaise.

This all changed in 1995 when, building on earlier work, Polchinski discovered D-branes, a class of extended solitonic objects upon which open strings end with Dirichlet boundary conditions. This proved to be the trigger for a second superstring revolution and was followed in quick succession by Witten's landmark discovery of M-theory and the web of string dualities connecting the five known 10-dimensional string theories and 11-dimensional supergravity that same year. Even more importantly, it was the direct antecedent of Maldacena's 1997 conjecture that quantum gravity (in the guise of Type IIB string theory on the 10-dimensional $AdS_5 \times S^5$) is *holographically* dual to a gauge theory (here, a maximally supersymmetric Yang–Mills theory living on the four-dimensional boundary of AdS_5). The impact that this duality has had on contemporary theoretical physics has been enormous, ranging from heavy ion physics and quantum criticality through emergent properties of spacetime[6] and the integrability structures of both string and gauge theories. Unfortunately, even after a decade of development, Maldacena's conjecture remains just that. So while a wealth of results have already been uncovered, there remains much

[4] Although it is worth pointing out that *noncritical* string theories can exist in any dimension ≤ 10.

[5] That an overwhelmingly large number of these solutions are supersymmetric, with no viable supersymmetry-breaking mechanism in sight, does not help much either.

[6] As is discussed in Chapter 9 by de Mello Koch and Murugan.

to be understood about what the AdS/CFT correspondence tells us about the nature of quantum gravity.

The developments outlined above form part of what might broadly be called the "covariant quantization" of gravity. Of course, in a field as diverse as quantum gravity, it is not the only program that has managed to make traction. A different approach to the problem is the "canonical quantization" of gravity. Based on the seminal 1967 work of DeWitt, this scheme utilizes the constrained Hamiltonian quantization, invented by Dirac in 1950 to quantize systems with gauge symmetries, to canonically quantize general relativity. A key characteristic of the canonical approach to quantum gravity is that it is *nonperturbative*. In contrast to perturbative formulations which require a choice to be made for a background spacetime metric from which to perturb, nonperturbative canonical methods have the advantage of being background-independent. This means that *all* aspects of space and time can, in principle, be determined from solutions of the theory.[7] In practice, however, this canonical approach was, for some time, stalled by the sheer intractability of the constraint (Wheeler–DeWitt) equation in the canonical variables of general relativity.

A major breakthrough came in the mid-1980s with Ashtekar's formulation of general relativity in terms of a new set of variables related to the holonomy group of the spacetime manifold. This in turn furnished a new basis for a nonperturbative quantization of general relativity in terms of Wilson loops. The result was the theory known as *loop quantum gravity*.[8] As one of the family of canonical quantum gravity theories, loop quantum gravity is both nonperturbative and manifestly background-independent. Among its major successes[9] are a nonperturbative quantization of 3-space geometry, a counting of the microstates of four-dimensional Schwarschild black holes and even a consistent truncation of its Hilbert space that suffices for questions of a cosmological nature to be addressed.[10] However, to pursue our analogy, on the battleground of quantum gravity, no single approach has yet proved faultless and the loop program (which includes LQG, spin-foam theories, loop quantum cosmology and, more recently, the group field theory of Oriti as outlined in Chapter 12) is no exception. Its critics point to, among other concerns, the lack of a consistent semiclassical limit that recovers general relativity and the necessarily *a posteriori* incorporation of the remaining interactions (as well as the matter content of the standard model).

[7] By contrast, in string theory for example, the dynamics of the string in spacetime should encode information about the spacetime metric so it would be preferable then that the metric not appear in the formulation of the theory. One solution to this problem is to find a viable nonperturbative formulation of the theory, as the AdS/CFT correspondence promises to provide.

[8] As described in Chapter 10 of this volume.

[9] Although, it must be said, these are not unequivocal.

[10] For an account of this so-called loop quantum cosmology, see Chapter 11 of this volume.

In addition to these two main research programs, the landscape of quantum gravity has been populated by a host of smaller, less developed, approaches that include Penrose's *twistor* program, *Regge calculus*, *Euclidean quantum gravity*, the *causal dynamical triangulations* of Ambjørn and Loll and Sorkin's *causal set theory*, each with its own fundamental tenet. The causal set program – introduced in this volume in Chapters 15 and 16, respectively – for example, is built on the principle that spacetime is fundamentally discrete with events related by a partial order that can be interpreted as an emergent causal structure.

This book has its origins in a (today, all too common) argument regarding the merits of string theory versus loop quantum gravity. After months of animated debate about quantization, symmetries, dimensions, background independence and innumerable other facets of the discussion, we realized that some of the questions we were meditating on might actually be useful to a broader community. These thoughts eventually crystallized in a wonderful workshop on the *Foundations of Space & Time* held at the Stellenbosch Institute for Advanced Study during August 2009 in honor of the 70th birthday of one of us (G. F. R. E.). The meeting brought together proponents of all the major programs in quantum gravity for a week of intense discussion and debate on the pros, cons, accomplishments, and shortcomings of each area.

By asking each speaker to be as open as possible about their own area and as curious as possible about each other's, we hoped to stimulate the kind of cross-field discussion that would make clear to everyone how far down the path to quantizing gravity we really are. The individual sessions were kept deliberately informal to facilitate such discussions. Interspersed among these were a number of focussed discussion sessions, with the most memorable of these revolving around two questions in particular. The first, "*Is spacetime fundamentally discrete or continuous?*," elicited several, varied responses with Lenny Susskind's (only partially tongue-in-cheek) "Yes!" being one of the most unexpected and (after some elaboration) interesting. For the second, open discussion, we posed the question: "*What do you want from a theory of quantum gravity?*," with the hopes of eliciting a wish-list of sorts from participants. The ensuing discussion was exactly what we expected; stimulating and insightful with answers ranging from testability at low energies to a complete understanding of the microscopic constituents of black holes.

In some ways we believe that we were enormously successful. In others not. On the one hand, language remains a significant problem in cross-field communication with only a very small set of researchers able to understand the technical nuances of other fields (and, consequently, appreciate some of the results therein). On the other hand, as disparate as they were in their approaches to the problem, almost everyone agreed that, even after all this time, the battle to reconcile quantum theory with gravity is far from over. These discussions, debates, and arguments were documented in the various contributions and synthesized into this volume. In this sense, this is arguably the most up-to-date account of where the field of quantum gravity currently stands. We hope that the reader will find reading this book as enjoyable as we did in putting it together.

2

A dialogue on the nature of gravity

T. PADMANABHAN

I describe the conceptual and mathematical basis of an approach which describes gravity as an emergent phenomenon. Combining the principle of equivalence and the principle of general covariance with known properties of local Rindler horizons, perceived by observers accelerated with respect to local inertial frames, one can provide a thermodynamic reinterpretation of the field equations describing gravity in any diffeomorphism-invariant theory. This fact, in turn, leads us to the possibility of deriving the field equations of gravity by maximizing a suitably defined entropy functional, without using the metric tensor as a dynamical variable. The approach synthesizes concepts from quantum theory, thermodynamics and gravity, leading to a fresh perspective on the nature of gravity. The description is presented here in the form of a dialogue, thereby addressing several frequently asked questions.

2.1 What is it all about?

Harold:[1] For quite some time now, you have been talking about 'gravity being an emergent phenomenon' and a 'thermodynamic perspective on gravity'. This is quite different from the conventional point of view in which gravity is a fundamental interaction and spacetime thermodynamics of, say, black holes is a particular result which can be derived in a specific context. Honestly, while I find your papers fascinating I am not clear about the broad picture you are trying to convey. Maybe

[1] Harold was a very useful creation originally due to Julian Schwinger [1] and stands for Hypothetically Alert Reader Of Limitless Dedication. In the present context, I think of Harold as Hypothetically Alert Relativist Open to Logical Discussions.

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

you could begin by clarifying what this is all about, before we plunge into the details? What is the roadmap, so to speak?

Me: To begin with, I will show you that the equations of motion describing gravity in *any diffeomorphism-invariant theory* can be given [2] a suggestive thermodynamic reinterpretation (Sections 2.2, 2.3). Second, taking a cue from this, I can formulate a variational principle for a suitably defined entropy functional – involving both gravity and matter – which will lead to the field equations of gravity [3,4] without varying the metric tensor as a dynamical variable (Section 2.4).

Harold: Suppose I have an action for gravity plus matter (in D dimensions)

$$A = \int d^D x \sqrt{-g} \left[L(R^{ab}_{cd}, g^{ab}) + L_{\rm matt}(g^{ab}, q_A) \right] \tag{2.1}$$

where L is any scalar built from metric and curvature and $L_{\rm matt}$ is the matter Lagrangian depending on the metric and some matter variables q_A. (I will assume L does not involve derivatives of curvature tensor, to simplify the discussion.) If I vary g^{ab} in the action I will get some equations of motion (see, e.g., [5,6]), say, $2E_{ab} = T_{ab}$ where E_{ab} is[2]

$$E_{ab} = P_a^{cde} R_{bcde} - 2\nabla^c \nabla^d P_{acdb} - \frac{1}{2} L g_{ab}; \quad P^{abcd} \equiv \frac{\partial L}{\partial R_{abcd}} \tag{2.2}$$

Now, you are telling me that (i) you can give a thermodynamic interpretation to the equation $2E_{ab} = T_{ab}$ just because it comes from a scalar Lagrangian and (ii) you can also derive it from an entropy maximization principle. I admit it is fascinating. But why should I take this approach as more fundamental, conceptually, than the good old way of just varying the total Lagrangian $L + L_{\rm matt}$ and getting $2E_{ab} = T_{ab}$? Why is it more than a curiosity?

Me: That brings me to the third aspect of the formulation which I will discuss towards the end (Section 2.5). In my approach, I can provide a natural explanation to several puzzling aspects of gravity and horizon thermodynamics, all of which have to be thought of as mere algebraic accidents in the conventional approach you mentioned. Let me give an analogy. In Newtonian gravity, the fact that inertial mass is equal to the gravitational mass is an algebraic accident without any fundamental explanation. But in a geometrical theory of gravity based on the principle of equivalence, this fact finds a natural explanation. Similarly, I think we can make progress by identifying key facts which have no explanation in the conventional approach and providing them a natural explanation from a different perspective. You will also see that this approach connects up several pieces of conventional theory in an elegant manner.

[2] The signature is $-+++$ and Latin letters cover spacetime indices while Greek letters run over space indices.

Harold: Your ideas also seem to be quite different from other works which describe gravity as an emergent phenomenon [7]. Can you explain *your* motivation?

Me: Yes. The original inspiration for my work, as for many others, comes from the old idea of Sakharov [8] which attempted to describe spacetime dynamics as akin to the theory of elasticity. There are two crucial differences between my approach and many other ones.

To begin with, I realized that the thermodynamic description transcends Einstein's general relativity and can incorporate a much wider class of theories – this was first pointed out in [9] and elaborated in several of my papers – while many other approaches concentrated on just Einstein's theory. In fact, many other approaches use techniques strongly linked to Einstein's theory – like, for example, the Raychaudhuri equation to study the rate of change of horizon area, which is difficult to generalize to theories in which the horizon entropy is *not* proportional to horizon area. I use more general techniques.

Second, I work at the level of action principle and its symmetries to a large extent so I have a handle on the off-shell structure of the theory; in fact, much of the thermodynamic interpretation in my approach is closely linked to the structure of action functional (like, e.g., the existence of surface term in action, holographic nature, etc.) for gravitational theories. This link is central to me while it is not taken into account in any other approach.

Harold: So essentially you are claiming that the thermodynamics of horizons is more central than the dynamics of the gravitational field while the conventional view is probably the other way around. Why do you stress the thermal aspects of horizons so much? Can you give a motivation?

Me: Because thermal phenomena is a window to microstructure! Let me explain. We know that the continuum description of a fluid, say, in terms of a set of dynamical variables like density ρ, velocity $\mathbf{v}$, etc. has a life of its own. At the same time, we also know that these dynamical variables and the description have no validity at a fundamental level where the matter is discrete. But one can actually *guess* the existence of microstructure without using any experimental proof for the molecular nature of the fluid, just from the fact that the fluid or a gas exhibits *thermal phenomena* involving temperature and transfer of heat energy. If the fluid is treated as a continuum and is described by $\rho(t, \mathbf{x})$, $\mathbf{v}(t, \mathbf{x})$, etc., all the way down, then it is *not* possible to explain the thermal phenomena in a natural manner. As first stressed by Boltzmann, the heat content of a fluid arises due to random motion of discrete microscopic structures which *must* exist in the fluid. These new degrees of freedom – which we now know are related to the actual molecules – make the fluid capable of storing energy internally and exchanging it with surroundings. So, given an apparently continuum phenomenon which exhibits temperature, Boltzmann could infer the existence of underlying discrete degrees of freedom.

Harold: I agree. But what does it lead to in the present context?
Me: The paradigm is: *If you can heat it, it has microstructure!* And you can heat up spacetimes by collapsing matter or even by just accelerating [10]. The horizons which arise in general relativity are endowed with temperatures [11] which shows that, at least in this context, some microscopic degrees of freedom are coming into play. So a thermodynamic description that links the standard description of gravity with the statistical mechanics of – as yet unknown – microscopic degrees of freedom must exist. It is in this sense that I consider gravity to be emergent.

Boltzmann's insight about the thermal behaviour has two other attractive features which are useful in our context. First, while the existence of the discrete degrees of freedom is vital in such an approach, the exact nature of the degrees of freedom is largely irrelevant. For example, whether we are dealing with argon molecules or helium molecules is largely irrelevant in the formulation of gas laws and such differences can be taken care of in terms of a few well-chosen numbers (like, e.g., the specific heat). This suggests that such a description will have a certain amount of robustness and independence as regards the precise nature of microscopic degrees of freedom.

Second, the entropy of the system arises due to our ignoring the microscopic degrees of freedom. Turning this around, one can expect the form of entropy functional to encode the essential aspects of microscopic degrees of freedom, even if we do not know what they are. If we can arrive at the appropriate form of entropy functional, in terms of some effective degrees of freedom, then we can expect it to provide the correct description.[3]
Harold: But most people working in quantum gravity will agree that there is some fundamental microstructure to spacetime ('atoms of spacetime') and the description of spacetime by metric is an approximate long-distance description. So why are you making a big deal? I don't see anything novel here.
Me: I will go farther than just saying there is microstructure and show you how to actually use the thermodynamic concepts to provide an emergent description of gravity – which no one else has attempted. If you think of the full theory of quantum gravity as analogous to statistical mechanics then I will provide the thermodynamic description of the same system.

As you know, thermodynamics was developed and used effectively decades before we knew anything about the molecular structure of matter or its statistical mechanics. Similarly, even without knowing the microstructure of spacetime or the full quantum theory of gravity, we can make a lot of progress with the

[3] Incidentally, this is why thermodynamics needed no modification due to either relativity or quantum theory. An equation like $TdS = dE + PdV$ will have universal applicability as long as effects of relativity or quantum theory are incorporated in the definition of $S(E, V)$ appropriately.

thermodynamic description of spacetime. The horizon thermodynamics, I will claim, provides [12] valuable insights about the nature of gravity totally independent of what 'the atoms of spacetime' may be. It is somewhat like being able to describe or work with gases or steam engines without knowing anything about the molecular structure of the gas or steam.

2.2 Local Rindler observers and entropy flow

Harold: All right. I hope all these will become clearer as we go along. Maybe I can suggest we plunge head-long into how *you* would like to describe gravity. Then I can raise the issues as we proceed.

Me: The overall structure of my approach is shown in Fig. 2.1. As you can see, I begin with the principle of equivalence which allows you to draw three key consequences. First, it tells you that – in the long wavelength limit – gravity has [13] a geometrical description in terms of the metric tensor g_{ab} and the effect of gravity on matter can be understood by using the laws of special relativity in the local inertial frames. Second, by writing Maxwell's equations in curved spacetime

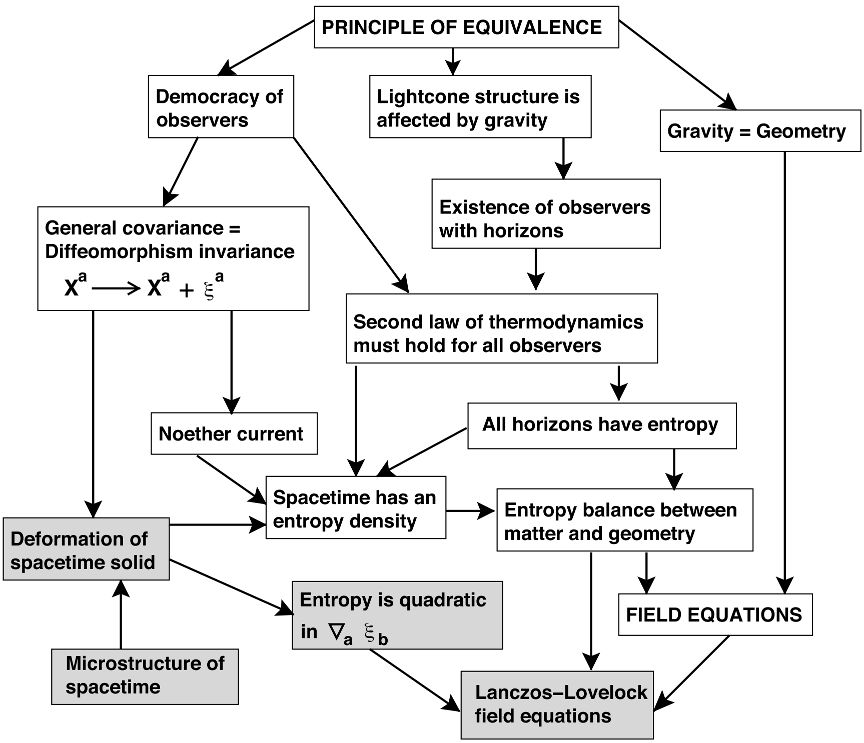


Figure 2.1 The broad picture

using minimal coupling, say, I can convince myself that the light cone structure of the spacetime – and hence the causal structure – will, in general, be affected by the gravitational field.

Harold: Well, that is one possible way of interpreting the principle of equivalence, though people might have other views. But once you have told me what you are assuming, viz., 'gravity = geometry' and 'light cones are affected by gravity', we can proceed further.

Me: Yes. My aim here will be not to nitpick over definitions but develop the physics in a consistent manner. In that spirit, I would draw one more conclusion from the fact that gravity can be described using a metric tensor. In flat spacetime, we can choose a special coordinate system with the global metric being η_{ab}; so if someone tells you that the metric is given by $g_{ab}(t,\mathbf{x})$ then you can always attribute the part $(g_{ab}-\eta_{ab})$ to the choice of non-inertial coordinates. We cannot do this in a curved spacetime. So it no longer makes sense to ask 'how much of g_{ab}' is due to our using non-inertial coordinates, and 'how much' is due to genuine gravity. Different observers in different states of motion might use different coordinates leading to different sets of 10 functions for $g_{ab}(t,\mathbf{x})$. Because we have no absolute reference metric it follows that no coordinate system or observer is special. The laws of physics should not select out any special class of observers.

Harold: This smacks of the principle of general covariance but essentially you are arguing [14] for democracy of all observers, which I grant you. Given all the philosophical controversies as to what 'general covariance' means, I agree this is a safer procedure. What next?

Me: Given the fact that all observers are equal and that light cones are affected by gravity, it follows that there will exist observers who do not have access to part of the spacetime because of the existence of horizons they perceive. This is a direct consequence of the fact that metric determines the paths of light rays and hence the causal structure. The classic example is the Rindler horizon in flat spacetime which is as effective in blocking information with respect to an accelerated observer as the Schwarzschild horizon at $r=2M$ is for an observer at $r>2M$.

Harold: Not so fast; I have several problems here. First, the conventional view is that black hole horizons are 'real horizons' while Rindler horizons are sort of fraudulent; you seem to club them together. Second, you seem to link horizons to observers rather than treat them as well-defined, geometrical, causal features of a spacetime.

Me: You are quite right. I treat all horizons at equal footing and claim that – for *my* purpose – all horizons are observer-dependent. This is because, the key property of horizons which I am concentrating on here is that they can block information. In that sense, the Rindler horizon does not block information for an inertial observer *just as* the Schwarzschild horizon does *not* block information for someone plunging into the black hole. So, for my purpose, there is no need to make artificial distinctions between a black hole horizon and a Rindler horizon. The state of motion of the observer is crucial in deciding the physical effects of a horizon *in all cases*.

It is, of course, true that one can give a geometric interpretation to, say, the black hole event horizon. I am not denying that. But that fact, as you will see, is quite irrelevant to the development of my approach.

Harold: I see that you not only demand democracy of observers but also democracy of horizons! You don't think, for example, that black hole horizons are anything special.

Me: Yes. I do believe in the democracy of horizons, as you put it. The attempts to provide a quantum gravitational interpretation of black holes, their entropy, etc. using *very special* approaches which are incapable of handling other horizons – like the issues in de Sitter [15], let alone Rindler – are interesting in a limited sort of way but may not get us anywhere ultimately.

Harold: I have another problem. You really haven't characterized what exactly you mean by a horizon for an observer. Of course, you cannot use any on-shell constructs since you are still developing your approach towards field equations. There are horizons and horizons in the literature – event, apparent, causal ...

Me: I will try to make clear what I *need* without again going into all sorts of definitions [2]. Choose any event $\mathcal{P}$ and introduce a local inertial frame (LIF) around it with Riemann normal coordinates $X^a = (T, \mathbf{X})$ such that $\mathcal{P}$ has the coordinates $X^a = 0$ in the LIF. Let k^a be a future directed null vector at $\mathcal{P}$ and we align the coordinates of LIF such that it lies in the $X - T$ plane at $\mathcal{P}$. Next transform from the LIF to local Rindler frame (LRF) coordinates x^a by accelerating along the X-axis with an acceleration κ by the usual transformation. The metric near the origin now reduces to the form

$$\begin{aligned} ds^2 &= -dT^2 + dX^2 + d\mathbf{x}_\perp^2 \\ &= -\kappa^2 x^2 dt^2 + dx^2 + d\mathbf{x}_\perp^2 = -2\kappa l\, dt^2 + \frac{dl^2}{2\kappa l} + d\mathbf{x}_\perp^2 \end{aligned} \tag{2.3}$$

where $T = x \sinh(\kappa t)$; $X = x \cosh(\kappa t)$, $l = (1/2)\kappa x^2$, and $(t, x, \mathbf{x}_\perp)$ or $(t, l, \mathbf{x}_\perp)$ are the coordinates of LRF (both these forms are useful in our discussion). Let ξ^a be the approximate Killing vector corresponding to translation in the Rindler time such that the vanishing of $\xi^a \xi_a \equiv -N^2$ characterizes the location of the local horizon $\mathcal{H}$ in LRF. As usual, we shall do all the computation on a timelike surface infinitesimally away from $\mathcal{H}$ with $N =$ constant, usually called a 'stretched horizon'. (It can be defined more formally using the orbits of ξ^a and the plane orthogonal to

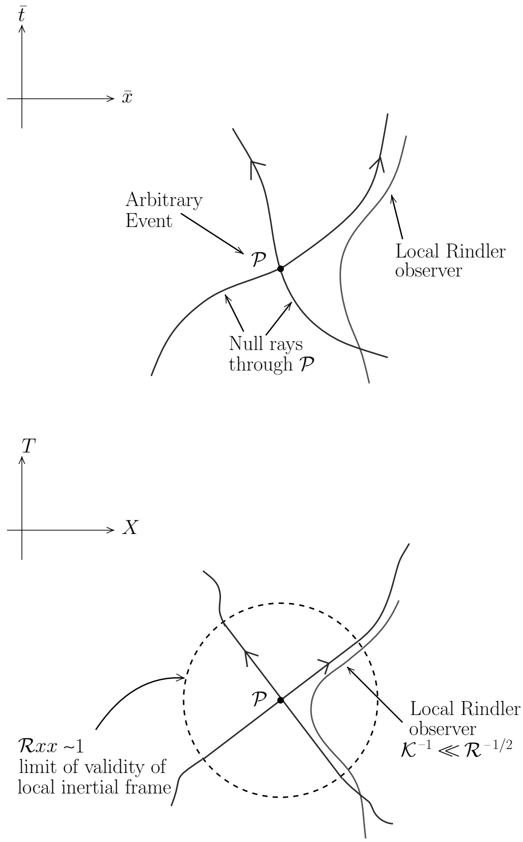


Figure 2.2 The top frame illustrates schematically the light rays near an event $\mathcal{P}$ in the $\bar{t} - \bar{x}$ plane of an arbitrary spacetime. The bottom frame shows the same neighbourhood of $\mathcal{P}$ in the locally inertial frame at $\mathcal{P}$ in Riemann normal coordinates (T, X). The light rays now become 45-degree lines and the trajectory of the local Rindler observer becomes a hyperbola very close to $T = \pm X$ lines which act as a local horizon to the Rindler observer.

the acceleration vector $a^i = \xi^b \nabla_b \xi^i$.) Let the timelike unit normal to the stretched horizon be r_a.

This LRF (with metric in Eq. (2.3)) and its local horizon $\mathcal{H}$ will exist within a region of size $L \ll \mathcal{R}^{-1/2}$ (where $\mathcal{R}$ is a typical component of curvature tensor of the background spacetime) as long as $\kappa^{-1} \ll \mathcal{R}^{-1/2}$. This condition can always be satisfied by taking a sufficiently large κ. This procedure introduces a class of uniformly accelerated observers who will perceive the null surface $T = \pm X$ as the local Rindler horizon $\mathcal{H}$. This is shown in Fig. 2.2.

Harold: I am with you so far. Essentially you are using the fact that you have two length scales in the problem at any event. First is the length scale $\mathcal{R}^{-1/2}$ associated

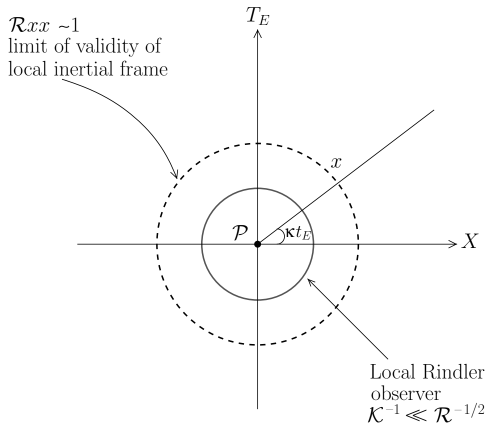


Figure 2.3 The region around $\mathcal{P}$ shown in Fig. 2.2 is represented in the Euclidean sector obtained by analytically continuing to imaginary values of T by $T_E = iT$. The horizons $T = \pm X$ collapse to the origin and the hyperbolic trajectory of the Rindler observer becomes a circle of radius κ^{-1} around the origin. The Rindler coordinates (t, x) become – on analytic continuation to $t_E = it$ – the polar coordinates $(r = x, \theta = \kappa t_E)$ near the origin.

with the curvature components of the background metric over which you have no control; second is the length scale κ^{-1} associated with the accelerated trajectory which you can choose as you please. So you can always ensure that $\kappa^{-1} \ll \mathcal{R}^{-1/2}$. In fact, I can see this clearly in the Euclidean sector in which the horizon maps to the origin (see Fig. 2.3). The locally flat frame in the Euclidean sector will exist in a region of radius $\mathcal{R}^{-1/2}$ while the trajectory of a uniformly accelerated observer will be a circle of radius κ^{-1}. You can always keep the latter inside the former. The metric in Eq. (2.3) is just the metric of the locally flat region in polar coordinates.
Me: Yes. In fact, I can choose a trajectory $x^i(\tau)$ such that its acceleration $a^j = u^i \nabla_i u^j$ (where u^i is the timelike 4-velocity) satisfies the condition $a^j a_j = \kappa^2$. In a suitably chosen LIF this trajectory will reduce to the standard hyperbola of a uniformly accelerated observer. It is using these LRFs that I define my horizons around any event. Further, the local temperature on the stretched horizon will be $\kappa/2\pi N$ so that $\beta_{\rm loc} = \beta N$ with $\beta \equiv \kappa/2\pi$.
Harold: Ha! The classical GR is fine but your 'horizon' is just a patch of null surface. Can you actually prove that local Rindler observers will perceive a temperature proportional to acceleration? The usual proofs of Unruh effect are [10] horribly global.
Me: Recall that everything we do is in a local region with $\kappa^{-1} \ll \mathcal{R}^{-1/2}$. Now if you have an accelerated detector with time-dependent, variable acceleration, say, then you will reproduce the standard Unruh effect *approximately* to the necessary

order of accuracy. This should be intuitively obvious but can be demonstrated [16]. Of course, in the Euclidean sector the Rindler observer's trajectory is a circle of radius κ^{-1} which can be made arbitrarily close to the origin. Suppose the observer's trajectory has the usual form $X = \kappa^{-1}\cosh\kappa t$; $T = \kappa^{-1}\sinh\kappa t$ which is maintained for a time interval $T \approx 2\pi/\kappa$. Then, the trajectory will complete a full circle *in the Euclidean sector* irrespective of what happens later! When we work in the limit of $\kappa \to \infty$, this becomes arbitrarily local in both space *and* time [17]. I am sure all these can be made more rigorous but this is the essential reason behind the local ideas working.

I also want to stress that once I finally reach my goal (of deriving the gravitational field equations from an entropy principle in Section 2.4) all these become irrelevant; they are essentially part of 'motivation'. So your possible misgivings regarding some of these details will not affect the final result.

2.3 Thermodynamic reinterpretation of the field equations

Harold: OK, so you have local Rindler observers crawling all over the spacetime with their local horizons. What next?

Me: It is now easy to see that all horizons must have entropy vis-à-vis the observers who perceive the horizons. If they do not, an observer can pour some hot tea with entropy across the horizon *à la Wheeler* [18], thereby violating the second law of thermodynamics in the region accessible to her and her friends who perceive the horizon $\mathcal{H}$. Such a violation of the second law of thermodynamics can be avoided only if we *demand* that the horizon should have an entropy which should increase when energy flows across it. If energy dE flows across a hot horizon with temperature T then $dE/T = dS$ should be the change in entropy of the horizon. We therefore conclude that all null surfaces which could locally act as one-way membranes should have an (observer-dependent) entropy associated with them.

Harold: Hold on. I understand from the reference you cite [18] that such a thought experiment might have had something to do with the initial realization of a black hole entropy which is proportional to the area [19]. But I am not sure how to interpret it precisely. For one thing, matter disappears into the horizon only after infinite time as perceived by the outside observer, even when you try to pour real tea into a real black hole. So what is all this talk about 'loss' of entropy?

Me: I don't think this is a real objection, though one often comes across this confusion. Note that, by the same argument, no black hole can ever form in finite time anywhere in the Universe and we should not be talking about any black hole physics. I believe this issue is well settled in chapter 33 of [13]. I recommend you read it!

If you really push me hard, I can wiggle out with the following argument. It does not take much time (certainly not infinite time!) for a cup of tea to reach a radial distance a few Planck lengths away from the horizon $r = 2M$. We have considerable evidence of a very different nature to suggest that the Planck length acts as a lower bound to the length scales that can be operationally defined and that no measurements can be ultra sharp at Planck scales [20]. So one cannot really talk about the location of the event horizon ignoring fluctuations of this order. So, from the point of view of sensible physics, I only need to get the cup of tea up to $r = 2M + L_P$ to talk about entropy loss.

Harold: You also seem to have quietly made entropy an observer-dependent quantity. *This is pretty drastic* and I need to understand it. Suppose, in a region around an event $\mathcal{P}$, there is some matter which is producing the curvature. I would normally have thought that this matter – say some amount of hot fluid – has a certain amount of entropy which is independent of who is measuring it. But you are claiming that an inertial observer and a Rindler observer will attribute different amounts of entropy to this matter. Is that correct?

Me: That's correct and maybe I should write a paper explaining this [21], but it really need not come as a surprise (also see [22]). We know that an inertial observer will attribute zero temperature and zero entropy to the inertial vacuum. But a Rindler observer will attribute a finite temperature and nonzero (formally divergent) entropy to the same vacuum state. So entropy is indeed an observer-dependent concept. When you do quantum field theory in curved spacetime, it is not only that particles become an observer-dependent notion, so do the temperature and entropy. This notion can be made more precise as follows.

Consider an excited state of a quantum field with energy δE above the ground state in an inertial spacetime. When you integrate out the unobservable modes for the Rindler observer, you will get a density matrix ρ_1 for this state and the corresponding entropy will be $S_1 = -\mathrm{Tr}\ (\rho_1 \ln \rho_1)$. The inertial vacuum state has the density matrix ρ_0 and the entropy $S_0 = -\mathrm{Tr}\ (\rho_0 \ln \rho_0)$. The difference $\delta S = S_1 - S_0$ is finite and represents the entropy attributed to this state by the Rindler observer. (This is finite though S_1 and S_0 can be divergent.) In the limit of $\kappa \to \infty$, in which we are working, we can actually compute it and show that

$$\delta S = \beta \delta E = \frac{2\pi}{\kappa} \delta E \tag{2.4}$$

To see this, note that if we write $\rho_1 = \rho_0 + \delta\rho$, then in the limit of $\kappa \to \infty$ we can concentrate on states for which $\delta\rho/\rho_0 \ll 1$. Then we have

$$\begin{aligned} -\delta S &= \mathrm{Tr}\ (\rho_1 \ln \rho_1) - \mathrm{Tr}\ (\rho_0 \ln \rho_0) \simeq \mathrm{Tr}\ (\delta\rho \ln \rho_0) \\ &= \mathrm{Tr}\ (\delta\rho(-\beta H_R)) = -\beta \mathrm{Tr}\ ((\rho_1 - \rho_0) H_R) \equiv -\beta \delta E \end{aligned} \tag{2.5}$$

where we have used the facts that Tr $\delta\rho \approx 0$ and $\rho_0 = Z^{-1}\exp(-\beta H_R)$, where H_R is the Hamiltonian for the system in the Rindler frame. The last line defines δE in terms of difference in expectation values of the Hamiltonian in the two states. (There are some subtleties in this derivation, especially regarding the assumption $\delta\rho/\rho_0 \ll 1$, but I will not get into it here [21].) This is the amount of entropy a Rindler observer would claim she has lost when the matter disappears into the horizon.

Harold: That is very curious. I would have thought that the expression for entropy of matter should consist of its energy δE and *its own* temperature T_{matter} rather than the horizon temperature. It looks like the matter somehow equilibrates to the horizon temperature so that $\delta S = \delta E/T_{\text{horizon}}$ gives the relevant entropy.

Me: Yes. This can be proved explicitly, for example, for the one-particle state [23] and here is a possible interpretation. You should think of the horizon as a system with some internal degrees of freedom and temperature *T as far as a Rindler observer is concerned*. So when you add an energy δE to it, the entropy change is $\delta S = (\delta E/T)$. All these are not *new* mysteries but only the manifestation of the old mystery, viz., a Rindler observer attributes a nonzero temperature to inertial vacuum. This temperature influences every other thermodynamic variable. I will come back to this point later on because it is quite important.

Harold: OK. Let us proceed. I also see where your insistence on democracy of observers comes in. You want to demand that the local Rindler observer has a right to expect the standard laws of physics to hold as much as any other observer, horizons notwithstanding.

Me: I am glad you brought this up. This was first pointed out in [14], in which we assert that *all observers have a right to describe physics using an effective theory based only on the variables they can access.* In the study of particle physics models, this concept forms the cornerstone of the renormalization group theory. To describe particle interactions at 10 GeV in the laboratory, we usually do not need to know what happens at 10^{14} GeV in theories which have predictive power. In the absence of such a principle, very high energy phenomena (which are unknown from direct experiments in the laboratory) will affect the low-energy phenomena which we are attempting to study.

In the context of a theory involving a nontrivial metric of spacetime, we need a similar principle to handle the fact that different observers will have access to different regions of a general spacetime. If a class of observers perceive a horizon, they should still be able to do physics using only the variables accessible to them without having to know what happens on the other side of the horizon.

This, in turn, implies that there should exist a mechanism which will encode the information in the region $\mathcal{V}$ which is inaccessible to a particular observer at the boundary $\partial\mathcal{V}$ of that region. Keep this in mind because I will show you later where this fits in with the holographic nature of action functionals.

Harold: Fine, we will get back to it. To get on with the story, you need to formulate some kind of entropy balance when matter flows across a local horizon. How do you propose to do it?

Me: Around any event in any spacetime we now have a local inertial frame and – by boosting along one of the axes with an acceleration κ – we have introduced a local Rindler observer who perceives a horizon with temperature proportional to κ. She will attribute a loss of entropy $\delta S = (2\pi/\kappa)\delta E$ when matter with an amount of energy δE gets close to the horizon (within a few Planck lengths, say). If ξ^a is the approximate Killing vector corresponding to translations in Rindler time, the appropriate energy–momentum density is $T^a_b \xi^b$. (It is the integral of $T^a_b \xi^b d\Sigma_a$ that gives the Rindler Hamiltonian H_R, which leads to evolution in Rindler time t and appears in the thermal density matrix $\rho = \exp -\beta H_R$.) The energy flux through a patch of stretched horizon with normal r_a will be $T_{ab}\xi^a r^b$ and the associated entropy flux will be $\beta_{\rm loc} T_{ab}\xi^a r^b$, where $\beta_{\rm loc}^{-1} = \beta^{-1}/N$ is the local temperature with N being the standard lapse function giving the redshift factor. (In conformity with Eq. (2.4), I am using the horizon temperature and not the matter temperature.) This entropy flux manifests as the entropy change of the locally perceived horizon. For all these to hold locally at every event there *must exist a spacetime entropy current* $\beta_{\rm loc} J^a$, built out of metric and its derivatives, such that $\beta_{\rm loc}(r_a J^a)$ gives the corresponding gravitational entropy flux. So we expect the relation

$$\beta_{\rm loc} r_a J^a = \beta_{\rm loc} T^{ab} r_a \xi_b \tag{2.6}$$

to hold at all events with some J^a, once we introduce a local Killing vector ξ^a and a local temperature giving $\beta_{\rm loc}$. Further, J^a must be conserved since we do not expect irreversible entropy production in the spacetime.

Harold: This sounds strange! Why should there exist a conserved current J^a, built from geometrical variables, at every event in some arbitrary spacetime, which will conveniently give you the entropy balance you require?

Me: It is actually not all that strange! Remember that we got into all these because of the democracy of the observers which, in turn, implies general covariance. The mathematical content of general covariance is captured by the diffeomorphism invariance of whatever theory is going to ultimately determine the dynamics of the spacetime. Because the diffeomorphism invariance of the theory forced us to treat all observers on an equal footing, the diffeomorphism invariance must also provide us with the conserved current J^a. And indeed it does, in the form of the Noether current [2]. Let me explain.

Consider a theory of gravity, obtained from a generally covariant action principle involving a gravitational Lagrangian $L(R^a_{bcd}, g^{ab})$ which is a scalar made from metric and curvature tensor. The total Lagrangian is the sum of L and the matter

Lagrangian L_m. The variation of the gravitational Lagrangian density generically leads to a surface term and hence can be expressed in the form

$$\delta(L\sqrt{-g}) = \sqrt{-g}\left(E_{ab}\delta g^{ab} + \nabla_a \delta v^a\right) \tag{2.7}$$

Under suitable boundary conditions the theory will lead to the field equation $2E_{ab} = T_{ab}$, where E_{ab} is given by Eq. (2.2) and T_{ab} is defined through the usual relation $(1/2)T_{ab}\sqrt{-g} = -(\delta A_m/\delta g^{ab})$. We also know that, for any Lagrangian L, the functional derivative E_{ab} satisfies the generalized off-shell Bianchi identity $\nabla_a E^{ab} = 0$.

Consider now the variations in δg_{ab} which arise through the diffeomorphism $x^a \to x^a + \xi^a$. In this case, $\delta(L\sqrt{-g}) = -\sqrt{-g}\nabla_a(L\xi^a)$, with $\delta g^{ab} = (\nabla^a \xi^b + \nabla^b \xi^a)$. Substituting these in Eq. (2.7) and using $\nabla_a E^{ab} = 0$, we obtain the conservation law $\nabla_a J^a = 0$, for the current

$$J^a \equiv \left(2E^{ab}\xi_b + L\xi^a + \delta_\xi v^a\right) \tag{2.8}$$

where $\delta_\xi v^a$ represents the boundary term which arises for the specific variation of the metric in the form $\delta g^{ab} = (\nabla^a \xi^b + \nabla^b \xi^a)$. It is also convenient to introduce the antisymmetric tensor J^{ab} by $J^a = \nabla_b J^{ab}$. Using the known expression for $\delta_\xi v^a$ in Eq. (2.8), it is possible to write an explicit expression for the current J^a for any diffeomorphism-invariant theory. For the general class of theories we are considering, the J^{ab} and J^a can be expressed [6] in the form

$$J^{ab} = 2P^{abcd}\nabla_c \xi_d - 4\xi_d\left(\nabla_c P^{abcd}\right) \tag{2.9}$$

$$J^a = -2\nabla_b\left(P^{adbc} + P^{acbd}\right)\nabla_c \xi_d + 2P^{abcd}\nabla_b\nabla_c\xi_d - 4\xi_d\nabla_b\nabla_c P^{abcd} \tag{2.10}$$

where $P_{abcd} \equiv (\partial L/\partial R^{abcd})$. These expressions simplify significantly at any event $\mathcal{P}$ where ξ^a behaves like an (approximate) Killing vector and satisfies the conditions

$$\nabla_{(a}\xi_{b)} = 0; \quad \nabla_a\nabla_b\xi_c = R_{cbad}\xi^d \tag{2.11}$$

(which a true Killing vector will satisfy everywhere). Then one can easily prove that $\delta_\xi v^a = 0$ at the event $\mathcal{P}$; the expression for Noether current simplifies considerably and is given by

$$J^a \equiv \left(2E^{ab}\xi_b + L\xi^a\right) \tag{2.12}$$

Harold: OK. So you now have a conserved current J^a and entropy current of matter. What do you do now?

Me: Recall that I argued, on very general grounds, that the relation in Eq. (2.6) *must* hold at all events. Remarkably enough, *the gravitational field equations of any diffeomorphism-invariant theory imply that this relation does hold!* To see this, let us now consider the form of $J^a(x)$ at any event $\mathcal{P}$ around which we have introduced the notion of a local Rindler horizon with ξ^a being the approximate Killing vector associated with the Rindler time translation invariance that satisfies two conditions in Eq. (2.11) at $\mathcal{P}$. Let r_a be the spacelike unit normal to the stretched horizon Σ, pointing in the direction of increasing N. We know that as $N \to 0$, the stretched horizon approaches the local horizon and Nr^i approaches ξ^i.

With this background, we compute J^a for the ξ^a introduced above in the neighbourhood of $\mathcal{P}$. Since it is an approximate Killing vector, satisfying Eq. (2.11), it follows that $\delta_\xi v = 0$ giving the current to be $J^a = \left(2E^{ab}\xi_b + L\xi^a\right)$. The product $r_a J^a$ for the vector r^a, which satisfies $\xi^a r_a = 0$ on the stretched horizon, becomes quite simple: $r_a J^a = 2E^{ab} r_a \xi_b$. This equation is valid around the local patch in which ξ^a is the approximate Killing vector. The quantity $\beta_{\text{loc}} r_a J^a$ (in this limit) is what we interpret as the local entropy flux density. On using the field equations $2E_{ab} = T_{ab}$, we immediately get

$$\beta_{\text{loc}} r_a J^a = 2E^{ab} r_a \xi_b = \beta_{\text{loc}} T^{ab} r_a \xi_b \tag{2.13}$$

which is exactly Eq. (2.6). This tells you that the validity of field equations in any diffeomorphism-invariant theory has a local, thermodynamic, interpretation. In the limit of $N \to 0$, this gives a *finite* result, $\beta \xi_a J^a = \beta T^{ab} \xi_a \xi_b$, as it should. Further, in this limit, ξ^i goes to $\kappa \lambda k^i$ where λ is the affine parameter associated with the null vector k^a we started with and all the reference to LRF goes away. It is clear that the properties of LRF are relevant conceptually to define the intermediate notions (local Killing vector, horizon temperature, etc.) but the essential result is independent of these notions. Just as we introduce a local inertial frame to decide how gravity couples to matter, we use local Rindler frames to interpret the physical content of the field equations.

Harold: That is cute! I also see why you can afford to be a bit cavalier about the LRF, etc.; ultimately, your interpretation is local at each event. The Noether current you use, of course, is the same that appears in the definition of Wald entropy [24]. But in the latter, it is used in an integral form while your approach seems to be completely local.

Me: This is true and I think the local approach is crucial for proper interpretation. Integrals over surfaces would require all sorts of special assumptions for everything to work out in an arbitrary spacetime. This is why I work in a local region around an arbitrary event with LIF, LRF, etc. with L in everything. Also note that the original definition of Wald entropy is an on-shell construct and requires you to evaluate an integral on a solution. The Noether current itself is an off-shell construct and that is what I need.

Incidentally, the Noether current relation can also be used to provide an alternative interpretation of the entropy balance along the following lines. A local Rindler observer, moving along the orbits of the Killing vector field ξ^a with four velocity $u^a = \xi^a/N$, will associate an energy $\delta E = u^a(T_{ab}\xi^b)dV_{\text{prop}}$ with a proper volume dV_{prop}. If this energy gets transferred across the horizon, the corresponding entropy transfer will be $\delta S_{\text{matter}} = \beta_{\text{loc}}\delta E$, where $\beta_{\text{loc}} = \beta N = (2\pi/\kappa)N$ is the local (redshifted) temperature of the horizon and N is the lapse function. Since $\beta_{\text{loc}}u^a = (\beta N)(\xi^a/N) = \beta\xi^a$, we find that

$$\delta S_{\text{matter}} = \beta\xi^a\xi^b T_{ab}\, dV_{\text{prop}} \tag{2.14}$$

As for gravitational entropy, since J^0 is the Noether charge density, $\delta S = \beta_{\text{loc}}u_a J^a dV_{\text{prop}}$ can be interpreted as the entropy associated with a volume dV_{prop} as measured by an observer with 4-velocity u^a. For observers moving on the orbits of the Killing vector ξ^a with $u^a = \xi^a/N$, we get

$$\delta S_{\text{grav}} = \beta N u_a J^a dV_{\text{prop}} = \beta[\xi_j\xi_a T^{aj} + L(\xi_j\xi^j)]\, dV_{\text{prop}} \tag{2.15}$$

As one approaches the horizon, $\xi^a\xi_a \to 0$, making the second term vanish and we get

$$\delta S_{\text{grav}} = \beta(\xi_j\xi_a T^{aj})\, dV_{\text{prop}} = \delta S_m \tag{2.16}$$

In the same limit ξ^j will become proportional to the original null vector k^j we started with. So this equation can again be thought of as an entropy balance condition.

Harold: So instead of thinking of field equations of gravity as $2E_{ab} = T_{ab}$, you want us to think of them as

$$[2E^{ab} - T^{ab}]k_a k_b = 0 \tag{2.17}$$

for all null vectors k^a. This is equivalent to $2E^{ab} - T^{ab} = \lambda g^{ab}$ with some constant λ. (I see that the constancy of λ follows from the conditions $\nabla_a E^{ab} = 0,\ \nabla_a T^{ab} = 0$.) Interpreting $2E_{ab}k^a k^b$ as some kind of gravitational entropy density and $T_{ab}k^a k^b$ as matter entropy in the local Rindler frame, you are providing a purely thermodynamical interpretation of the field equations of any diffeomorphism-invariant theory of gravity. Right?

Me: Yes. But note that Eq. (2.17) is *not* quite the same as the standard equation $2E^{ab} = T^{ab}$ because Eq. (2.17) has an extra symmetry which standard gravitational field equations do not have: this equation is invariant under the shift $T^{ab} \to T^{ab} + \mu g^{ab}$ with some constant μ. (This symmetry has important implications for the cosmological constant problem which we will discuss later.) While the properties of LRF are relevant conceptually to define the intermediate notions (local Killing vector, horizon temperature, etc.), the essential result is independent of these notions.

Harold: Fine. I like the fact that *just as we introduce local inertial frames to decide how gravity couples to matter, we use local Rindler frames to interpret the physical content of the field equations.* But you only needed the part of J^a given by $2E^a_b\xi^b$ for your analysis, right? The other two terms in Eq. (2.8) are not needed at all. So maybe you don't have to use all of the Noether current.

Me: This is quite true. In fact, one can give $2E^a_b\xi^b$ an interesting interpretation. Suppose there are some microscopic degrees of freedom in spacetime, just as there are atoms in a solid. If you make the solid undergo an elastic deformation $x^\alpha \to x^\alpha + \xi^\alpha(x)$, the physics can be formulated in terms of the displacement field $\xi^\alpha(x)$ and one can ask how thermodynamic potentials like entropy change under such displacement. Similarly, in the case of spacetime, we should think of

$$\delta S_{\rm grav} = \beta_{\rm loc}(2E^a_b)u_a\delta x^b \tag{2.18}$$

as the change in the gravitational entropy under the 'deformation' of the spacetime $x^a \to x^a + \delta x^a$ as measured by the Rindler observer with velocity u^a. One can show that this interpretation is consistent with all that we know about horizon thermodynamics. So the left-hand side of the gravitational field equation $(2E^a_b)$ actually gives the response of the spacetime entropy to the deformations.

Harold: It certainly matches with the previous results. Since $\beta_{\rm loc}u_a = \beta\xi_a$, you will get the entropy density to be proportional to $2E_{ab}k^ak^b$ on the horizon. Does it make sense?

Me: As I will show you soon, it makes a lot of sense!

Harold: But can't you now reverse the argument and claim that you can derive the field equations of the theory from the purely thermodynamic point of view of the entropy balance?

Me: That would be lovely and very tempting, but I don't think so. Such a 'reverse-engineering' faces some conceptual hurdles; the mathematics we will go through trivially but not the logic [2]. Let me clarify the issues involved.

The key point is the following: if we have a justification for interpreting the expression $\beta_{\rm loc}(r_aJ^a)$ as entropy current, *independent of the field equations*, then – *and only then* – can we invert the logic and obtain the field equations from the thermodynamic identity. However, in the absence of field equations, J^a is just a Noether current. It can be interpreted as an entropy current *if and only if* field equations are assumed to hold; it is in this on-shell context that Wald [24] showed that it is entropy. So we have no *independent* justification for demanding $\beta_{\rm loc}r_aJ^a$ should be equal to matter entropy flux. Until we come up with such a justification – without using field equations – we can prove that 'field equations imply local entropy balance at local

horizons' but not 'local entropy balance at local horizons imply field equations'. The issue at stake is not mathematics but logic. As a simple example, consider the Noether current in Einstein's theory for a Killing vector ξ^a, which is proportional to $R^a_b \xi^b$. No one would have thought of this expression as entropy density independent of field equations. It is only by studying physical processes involving black holes, say, *and using field equations* that one can give such a meaning.
Harold: OK. I have one more worry. At this stage, you have not chosen any specific theory of gravity at all, right? So this thermodynamic entropy balance seems to be very general and some people might even say it is *too general*. What is your take on this?
Me: It is true that at this stage I have not specified what kind of theory of gravity we are dealing with. The field equation – whatever the theory may be, as long as it obeys the principle of equivalence and diffeomorphism invariance – always has an interpretation in terms of local entropy balance. (The idea also works when L_{grav} depends on the derivatives of the curvature tensor but I will not discuss this case, for the sake of simplicity.) *Different theories of gravity are characterized by different forms of entropy density just as different physical systems are characterized by different forms of entropy functionals.* I think this is completely in harmony with the thermodynamic spirit. Thermodynamics applies to any system; if you want to describe a *particular* system, you need to specify its entropy functional or some other thermodynamic potential. So what the development so far is telling us is that we need to put in some more extra physical input into the theory to find the field equations describing the theory.

2.4 Field equations from a new variational principle

Harold: Fine. The above results imply that the field equations arising from any generally covariant action can be given a thermodynamic interpretation; that is, you assumed the validity of the field equations and derived the local entropy balance. Your real aim, however, is to obtain the field equations from a dynamical principle rather than *assume* the field equations. How do you propose to do that?
Me: To begin with, I want to paraphrase the above results in a slightly different manner which is probably more useful for the task we want to undertake.

Note that, instead of dropping matter across the horizon, I could have equally well considered a virtual, infinitesimal (Planck scale), displacement of the $\mathcal{H}$ normal to itself engulfing some matter. We only need to consider infinitesimal displacements because the entropy of the matter is not 'lost' until it crosses the horizon; that is, until the matter is at an infinitesimal distance (a few Planck lengths) from the horizon. All the relevant physical processes take place at a region very close to the horizon and hence an infinitesimal displacement of $\mathcal{H}$ normal to itself will engulf some matter. Some entropy will again be lost to the outside observers unless displacing a piece of local Rindler horizon costs some entropy.

So we expect the entropy balance condition derived earlier to ensure this and indeed it does. An infinitesimal displacement of a local patch of the stretched horizon in the direction of r_a, by an infinitesimal proper distance ε, will change the proper volume by $dV_{\rm prop} = \varepsilon\sqrt{\sigma}\,d^{D-2}x$, where σ_{ab} is the metric in the transverse space. The flux of energy through the surface will be $T^a_b\xi^b r_a$ and the corresponding entropy flux can be obtained by multiplying the energy flux by $\beta_{\rm loc}$. Hence the 'loss' of matter entropy to the outside observer when the virtual displacement of the horizon swallows some hot tea is $\delta S_m = \beta_{\rm loc}\delta E = \beta_{\rm loc}T^{aj}\xi_a r_j dV_{\rm prop}$. To find the change in the gravitational entropy, we again use the Noether current J^a corresponding to the local Killing vector ξ^a. Multiplying by r^a and $\beta_{\rm loc} = \beta N$, we get

$$\beta_{\rm loc} r_a J^a = \beta_{\rm loc}\xi_a r_a T^{ab} + \beta N(r_a\xi^a)L \tag{2.19}$$

As the stretched horizon approaches the true horizon, we know that $Nr^a \to \xi^a$ and $\beta\xi^a\xi_a L \to 0$, making the last term vanish. So

$$\delta S_{\rm grav} \equiv \beta\xi_a J^a dV_{\rm prop} = \beta T^{aj}\xi_a\xi_j dV_{\rm prop} = \delta S_m \tag{2.20}$$

showing again the validity of local entropy balance.

Harold: It appears to me that this is similar to switching from a passive point of view to an active point of view. Instead of letting a cup of tea fall into the horizon, you are making a virtual displacement of the horizon surface to engulf the tea, which is infinitesimally close to the horizon. But in the process, you have introduced the notion of virtual displacement of horizons and for the theory to be consistent, this displacement of these surface degrees of freedom should cost you some entropy. Right?

Me: Yes. If gravity is an emergent, long-wavelength, phenomenon like elasticity then the diffeomorphism $x^a \to x^a + \xi^a$ is analogous to the elastic deformations of the 'spacetime solid' [25]. It then makes sense to demand that the entropy density should be a functional of ξ^a and their derivatives $\nabla_b\xi^a$. By constraining the functional form of this entropy density, we can choose the field equations of gravity. Recall that thermodynamics relies entirely on the form of the entropy functional to make predictions. If we constrain the form of the entropy, we constrain the theory.

So the next step is to assume a suitable form of entropy functional for gravity $S_{\rm grav}$ in terms of the normal to the null surface. Then it seems natural to demand that the dynamics should follow from the extremum prescription $\delta[S_{\rm grav} + S_{\rm matter}] = 0$ for *all null surfaces in the spacetime* where $S_{\rm matter}$ is the matter entropy.

Harold: What do we take for $S_{\rm grav}$ and $S_{\rm matter}$?
Me: The form of $S_{\rm matter}$ is easy to ascertain from the previous discussion. If T_{ab} is the matter energy–momentum tensor in a general $D\,(\geq 4)$-dimensional spacetime then an expression for matter entropy *relevant for our purpose* can be taken to be

$$S_{\rm matter} = \int_{\mathcal{V}} d^D x \sqrt{-g} T_{ab} n^a n^b \tag{2.21}$$

where n^a is a null vector field. From our Eq. (2.14) we see that the entropy density associated with proper 3-volume is $\beta(T_{ab}\xi^a\xi^b)dV_{\rm prop}$, where – on the horizon – the vector ξ^a becomes proportional to a null vector n^a. If we now use the Rindler coordinates in Eq. (2.3) in which $\sqrt{-g}=1$ and interpret the factor β as arising from an integration of dt in the range $(0,\beta)$, we find that the entropy density associated with a proper 4-volume is $(T_{ab}n^a n^b)$. This suggests treating Eq. (2.21) as the matter entropy. For example, if T_{ab} is due to an ideal fluid at rest in the LIF then $T_{ab}n^a n^b$ will contribute $(\rho+P)$, which – by the Gibbs–Duhem relation – is just $T_{\rm local}s$ where s is the entropy density and $T_{\rm local}^{-1}=\beta N$ is the properly redshifted temperature with $\beta=2\pi/\kappa$ being the periodicity of the Euclidean time coordinate. Then

$$\begin{aligned}\int dS &= \int \sqrt{h} d^3x\; s = \int \sqrt{h} d^3x \beta_{\rm loc}(\rho+P) = \int \sqrt{h} N d^3x \beta(\rho+P)\\ &= \int_0^\beta dt \int d^3x \sqrt{-g}\, T^{ab} n_a n_b\end{aligned} \tag{2.22}$$

which matches with Eq. (2.21) in the appropriate limit.
Harold: Well, that may be all right for an ideal fluid. But for a general source, like say the electromagnetic field (which will act as a source in Reissner–Nordstrom metric) I don't even know how to define entropy. But I am willing to accept Eq. (2.21) as a definition.
Me: Actually, it is better than that. We *do* have the notion of *energy* flux across a surface with normal r^a being $T_{ab}\xi^b r^a$ which holds for *any* source T^{ab}. Given some energy flux δE in the Rindler frame, there is an associated entropy flux loss $\delta S=\beta_{\rm hor}\delta E$ as given by Eq. (2.4). You may think that an ordered field has no temperature or entropy, but a Rindler observer will say something different. For any state, she will have a corresponding density matrix ρ and an entropy $-Tr(\rho\ln\rho)$; after all, she will attribute entropy even to the vacuum state. It is *this* entropy which is given by Eq. (2.14) and Eq. (2.21).
Harold: Interesting. There is also this time integration which you limit to the range $(0,\beta)$ in Eq. (2.22). This is fine in the Euclidean sector and maybe you can rotate back to the Lorentzian sector but it makes me a little uncomfortable.

Me: Well, I again have to the invoke the local nature of the argument which, as we discussed, is obvious in the Euclidean sector but the concept of causality, loss of information, etc. is obvious in the Lorentzian sector in which I have light cones and null surfaces. So I do have to switch back and forth. Maybe there is a better way of formulating this which I have not yet figured out; but for our purpose, you can even think of all integrals being done in the Euclidean sector – if you are happier with that.

Harold: Fine. What about $S_{\rm grav}$?

Me: For this, I will first describe the simplest possible choice and will then consider a more general expression. The simplest choice is to postulate $S_{\rm grav}$ to be a quadratic expression [3] in the derivatives of the normal:

$$S_{\rm grav} = -4\int_{\mathcal{V}} d^D x\sqrt{-g}\, P_{ab}^{cd}\nabla_c n^a \nabla_d n^b \tag{2.23}$$

where the explicit form of P_{ab}^{cd} is ascertained below. The expression for the total entropy now becomes

$$S[n^a] = -\int_{\mathcal{V}} d^D x\sqrt{-g}\left(4P_{ab}^{cd}\nabla_c n^a \nabla_d n^b - T_{ab}n^a n^b\right) \tag{2.24}$$

If you want, you can forget everything we said so far and start with this expression as defining our theory!

Harold: I suppose this is your variational principle and you will now extremize S with respect to n_a?

Me: Yes, but I want to first explain the crucial conceptual difference between the extremum principle introduced here and the conventional one. Usually, given a set of dynamical variables n_a and a functional $S[n_a]$, the extremum principle will give a set of equations for the dynamical variable n_a. Here the situation is completely different. We expect the variational principle to hold for *all* null vectors n^a, thereby leading to a condition on the *background metric*. (Of course, one can specify any null vector $n^a(x)$ by giving its components $f^A(x) \equiv n^a e_a^A$ with respect to a fixed set of basis vectors e_a^A with $e_A^b e_b^B = \delta_A^B$, etc. so that $n^a = f^A e_A^a$. So the class of all null vectors can be mapped to the scalar functions f^A with the condition $f_A f^A = 0$.) Obviously, the functional in Eq. (2.24) must be rather special to accomplish this and one needs to impose restrictions on P_{ab}^{cd} (and T_{ab}, though that condition turns out to be trivial) to achieve this.

It turns out – as we shall see below – that two conditions are sufficient to ensure this. First, the tensor P_{abcd} should have the same algebraic symmetries as the Riemann tensor R_{abcd} of the D-dimensional spacetime. This condition can be ensured if we define P_a^{bcd} as

$$P_a^{bcd} = \frac{\partial L}{\partial R^a_{bcd}} \tag{2.25}$$

where $L = L(R^a_{bcd}, g^{ik})$ is some scalar. Second, I will postulate the condition:

$$\nabla_a P^{abcd} = 0 \tag{2.26}$$

as well as $\nabla_a T^{ab} = 0$, which is anyway satisfied by any matter energy–momentum tensor.

Harold: Can you give some motivation for these conditions?

Me: As regards Eq. (2.25), the motivation will become clearer later on. Basically, I will show that this approach leads to the same field equations as the one with L as gravitational Lagrangian in the conventional approach. (That is why I have used the symbol L for this scalar!)

One possible motivation for Eq. (2.26) arises from the fact that it will ensure the field equations do not contain any derivative higher than second order of the metric. Another possible interpretation arises from the analogy introduced earlier. If you think of n^a as analogous to the deformation field in elasticity, then, in the theory of elasticity [26] one usually postulates the form of the thermodynamic potentials which are quadratic in first derivatives of n_a. The coefficients of this term will be the elastic constants. Here the coefficients are P^{abcd} and you may want to think of Eq. (2.26) as saying the ‘elastic constants of spacetime solid’ are actually ‘constants’. But nothing depends on this picture. In fact, I will show you later how this condition in Eq. (2.26) can be relaxed.

Harold: Interesting. You claim extremizing Eq. (2.24) in this context with respect to all n^a leads to an equation constraining the background metric. If so, this is a peculiar variational principle.

Me: Let me show you how this arises. Varying the normal vector field n^a after adding a Lagrange multiplier function $\lambda(x)$ for imposing the condition $n_a \delta n^a = 0$, we get

$$-\delta S = 2\int_{\mathcal{V}} d^D x\sqrt{-g}\left(4P_{ab}^{cd}\nabla_c n^a\left(\nabla_d \delta n^b\right) - T_{ab}n^a\delta n^b - \lambda(x) g_{ab}n^a\delta n^b\right) \tag{2.27}$$

where we have used the symmetries of P_{ab}^{cd} and T_{ab}. An integration by parts and the condition $\nabla_d P_{ab}^{cd} = 0$, leads to

$$-\delta S = 2\int_{\mathcal{V}} d^D x\sqrt{-g}\left[-4P_{ab}^{cd}\left(\nabla_d\nabla_c n^a\right) - (T_{ab} + \lambda g_{ab})n^a\right]\delta n^b$$
$$+8\int_{\partial\mathcal{V}} d^{D-1}x\sqrt{h}\left[k_d P_{ab}^{cd}\left(\nabla_c n^a\right)\right]\delta n^b \tag{2.28}$$

where k^a is the D-vector field normal to the boundary $\partial\mathcal{V}$ and h is the determinant of the intrinsic metric on $\partial\mathcal{V}$. As usual, in order for the variational principle to be well defined, we require that the variation δn^a of the vector field should vanish on the boundary. The second term in Eq. (2.28) therefore vanishes, and the condition that $S[n^a]$ be an extremum for arbitrary variations of n^a then becomes

$$2P_{ab}^{cd}\left(\nabla_c\nabla_d - \nabla_d\nabla_c\right)n^a - (T_{ab} + \lambda g_{ab})n^a = 0 \tag{2.29}$$

where we used the antisymmetry of P_{ab}^{cd} in its upper two indices to write the first term. The definition of the Riemann tensor in terms of the commutator of covariant derivatives reduces the above expression to

$$\left(2P_b^{ijk}R^a_{ijk} - T^a_b + \lambda\delta^a_b\right)n_a = 0 \tag{2.30}$$

and we see that the equations of motion *do not contain* derivatives with respect to n^a which is, of course, the crucial point. This peculiar feature arose because of the symmetry requirements we imposed on the tensor P_{ab}^{cd}. We require that the condition in Eq. (2.30) holds for *arbitrary* vector fields n^a. One can easily show [3] that this requires

$$16\pi\left[P_b^{ijk}R^a_{ijk} - \frac{1}{2}\delta^a_b L\right] = 8\pi T^a_b + \Lambda\delta^a_b \tag{2.31}$$

Comparison with Eq. (2.2) shows that these are precisely the field equations for gravity (with a cosmological constant arising as an undetermined integration constant; more about this later) in a theory with Lagrangian L when Eq. (2.24) is satisfied. That is, we have $2E_{ab} = T_{ab} + \lambda g_{ab}$ with

$$E_{ab} = P_a^{cde}R_{bcde} - \frac{1}{2}Lg_{ab}; \quad P^{abcd} \equiv \frac{\partial L}{\partial R_{abcd}} \tag{2.32}$$

The crucial difference between Eq. (2.2) and Eq. (2.32) is that, the E_{ab} in Eq. (2.32) contains no derivatives of the metric higher than second order thereby leading to field equations which are second order in the metric. In contrast, Eq. (2.2) can contain up to fourth-order derivatives of the metric.

Harold: Let me get this straight. Suppose I start with a total Lagrangian $L(R_{abcd}, g_{ab}) + L_{\text{matt}}$, and define a P^{abcd} by Eq. (2.25) ensuring it satisfies Eq. (2.26). Then I get certain field equations by varying the metric. You have just proved that I will get the *same* field equations (but with a cosmological constant) if I start with the expression in Eq. (2.24), maximize it with respect to n^a and demand that it holds for all n^a. The maths is clear but I have several doubts. To begin with, why does the maths work out?!

Me: I will let you into the secret by doing it differently. Note that, using the constraints on P^{abcd} I can prove the identity

$$\begin{aligned}4P_{ab}^{cd}\nabla_c n^a \nabla_d n^b &= 4\nabla_c[P_{ab}^{cd} n^a \nabla_d n^b] - 4n^a P_{ab}^{cd}\nabla_c \nabla_d n^b \\ &= 4\nabla_c[P_{ab}^{cd} n^a \nabla_d n^b] - 2n^a P_{ab}^{cd}\nabla_{[c}\nabla_{d]} n^b \\ &= 4\nabla_c[P_{ab}^{cd} n^a \nabla_d n^b] - 2n^a P_{ab}^{cd} R^b_{icd} n^i \\ &= 4\nabla_c[P_{ab}^{cd} n^a \nabla_d n^b] + 2n^a E_{ai} n^i \end{aligned} \tag{2.33}$$

where the first line uses Eq. (2.26), the second line uses the antisymmetry of P_{ab}^{cd} in c and d, the third line uses the standard identity for a commutator of covariant derivatives and the last line is based on Eq. (2.2) when $n_a n^a = 0$ and Eq. (2.26) hold. Using this in the expression for S in Eq. (2.24) and integrating the 4-divergence term, I can write

$$S[n^a] = -\int_{\partial\mathcal{V}} d^{D-1}x k_c \sqrt{h}(4P_{ab}^{cd} n^a \nabla_d n^b) - \int_{\mathcal{V}} d^D x \sqrt{-g}\left[(2E_{ab} - T_{ab})n^a n^b\right] \tag{2.34}$$

So, when I consider variations ignoring the surface term I am effectively varying $(2E_{ab} - T_{ab})n^a n^b$ with respect to n_a and demanding that it holds for all n_a. That should explain to you why it leads to $(2E_{ab} = T_{ab})$ except for a cosmological constant.

Harold: Ha! I see it. Of course, there is an ambiguity of adding a term of the form $\lambda(x)g_{ab}$ in the integrand of the second term in Eq. (2.34), leading to the final equation $(2E_{ab} = T_{ab} + \lambda(x)g_{ab})$ but the Bianchi identity $\nabla_a E^{ab} = 0$ along with $\nabla_a T^{ab} = 0$ will make $\lambda(x)$ actually a constant.

Me: Yes. Remember that. We will discuss the cosmological constant issue separately in the end.

Harold: I see that it also connects up with your previous use of $2E_{ab}n^a n^b$ as some kind of gravitational entropy density. Your expression for gravitational entropy is actually

$$
\begin{aligned}
S_{\text{grav}}[n^a] &= -\int_{\mathcal{V}} d^D x \sqrt{-g}\, 4P_{ab}^{cd} \nabla_c n^a \nabla_d n^b \\
&= -\int_{\partial\mathcal{V}} d^{D-1} x k_c \sqrt{h} (4P_{ab}^{cd} n^a \nabla_d n^b) - \int_{\mathcal{V}} d^D x \sqrt{-g} (2E_{ab} n^a n^b)
\end{aligned}
\tag{2.35}
$$

Written in this form you have a bulk contribution (proportional to our old friend $2E_{ab}n^a n^b$) and a surface contribution. When equations of motion hold, the bulk also gets a contribution from matter which cancels it out, leaving the entropy of a region $\mathcal{V}$ to reside in its boundary $\partial\mathcal{V}$.

Me: Yes. I need to think more about this.

Harold: Also, arising out of your letting me into the trick, I realize that I can now find an S for any theory, even if Eq. (2.25) does not hold. You just have to reverse-engineer it starting from $(2E_{ab} - T_{ab})n^a n^b$ as the entropy density and using the expression in Eq. (2.2) for E_{ab}, right? So why do you insist on Eq. (2.25)?

Me: You are right, of course. If you start with $(2E_{ab} - T_{ab})n^a n^b$ as the entropy density (see eq. (14) of the first paper in [2]) and work backwards you will get for S_{grav} the expression

$$
\begin{aligned}
S_{\text{grav}} = -4 \int_V d^D x \sqrt{-g} \Big(& P^{abcd} \nabla_c n_a \nabla_d n_b + (\nabla_d P^{abcd}) n_b \nabla_c n_a \\
& + (\nabla_c \nabla_d P^{abcd}) n_a n_b \Big)
\end{aligned}
\tag{2.36}
$$

Varying this with respect to n^a will then lead to the correct equations of motion and – incidentally – the same surface term.

While one could indeed work with this more general expression, there are four reasons to prefer the imposition of the condition in Eq. (2.25). First, it is clear from Eq. (2.2) that when L depends on the curvature tensor and the metric, E_{ab} can depend on up to the fourth derivative of the metric if Eq. (2.25) is not satisfied. But when we impose Eq. (2.25) then we are led to field equations which have, at most, second derivatives of the metric tensor – which is a desirable feature. Second, as we shall see below, with that condition we can actually determine the form of L; it turns out that in $D = 4$, it uniquely selects Einstein's theory, which is probably a nice feature. In higher dimensions, it picks out a very geometrical extension of Einstein's theory in the form of Lanczos–Lovelock theories. Third, it is difficult to imagine why the terms in Eq. (2.36) should occur with very specific coefficients. In fact, it is not clear why we cannot have derivatives of R_{abcd} in L, if the derivatives of P_{abcd} can occur in the expression for entropy. Finally, if we take the idea of elastic constants being constants, then one is led to Eq. (2.25). None of these rigorously exclude the possibility in Eq. (2.36), and in fact this model has been explored recently [27].

Harold: So far we have not fixed P^{abcd} so we have not fixed the theory. How does Eq. (2.25) allow you to do this?

Me: In a complete theory, the explicit form of P^{abcd} will be determined by the long-wavelength limit of the microscopic theory just as the elastic constants can – in principle – be determined from the microscopic theory of the lattice. In the absence of such a theory, we need to determine P^{abcd} by general considerations. Essentially we need to determine scalar L built from curvature tensor and the metric which satisfies the constraint $\nabla_a(\partial L/\partial R_{abcd})=0$. This problem can be solved completely, and the result is the Lagrangian of a Lanczos–Lovelock theory. Such an L can be written as a sum of terms, each involving products of curvature tensors with the m-th term being a product of m curvature tensors, leading to

$$L=\sum_{m=1}^{K} c_m L_{(m)};\; L_{(m)}=\frac{1}{16\pi}2^{-m}\delta^{a_1a_2\ldots a_{2m}}_{b_1b_2\ldots b_{2m}} R^{b_1b_2}_{a_1a_2}\cdots R^{b_{2m-1}b_{2m}}_{a_{2m-1}a_{2m}} \tag{2.37}$$

where the c_m are arbitrary constants and $L_{(m)}$ is the mth-order Lanczos–Lovelock Lagrangian. The $m=1$ term is proportional to $\delta^{ab}_{cd}R^{cd}_{ab}\propto R$ and leads to Einstein's theory. It is conventional to take $c_1=1$, so that the $L_{(1)}$ reduces to $R/16\pi$. The normalizations for $m>1$ are somewhat arbitrary for individual $L_{(m)}$ since the c_ms are unspecified at this stage. The $m=2$ term gives rise to what is known as Gauss–Bonnet theory. Because of the determinant tensor, it is obvious that in any given dimension D we can only have K terms where $2K\le D$. It follows that, if $D=4$, then only the $m=1,2$ are nonzero. Of these, the Gauss–Bonnet term (corresponding to $m=2$) gives, on variation of the action, a vanishing bulk contribution in $D=4$. (In dimensions $D=5$ to 8, one can have both the Einstein–Hilbert term and the Gauss–Bonnet term and so on.) Equivalently, the P^{abcd} can be expressed as a series in the powers of derivatives of the metric as

$$P^{abcd}(g_{ij},R_{ijkl})=c_1\overset{(1)}{P}{}^{abcd}(g_{ij})+c_2\overset{(2)}{P}{}^{abcd}(g_{ij},R_{ijkl})+\cdots \tag{2.38}$$

where $c_1,c_2,\ldots$ are coupling constants. The lowest-order term depends only on the metric with no derivatives. The next term depends (in addition to the metric) linearly on the curvature tensor, and the next one will be quadratic in curvature, etc.

Let us take a closer look at the structure which is emerging. The lowest-order term in Eq. (2.38) (which leads to Einstein's theory) is

$$\overset{(1)}{P}{}^{ab}_{cd}=\frac{1}{16\pi}\frac{1}{2}\delta^{ab}_{cd}=\frac{1}{32\pi}(\delta^a_c\delta^b_d-\delta^a_d\delta^b_c) \tag{2.39}$$

To the lowest order, when we use Eq. (2.39) for P_b^{ijk}, Eq. (2.31) reduces to Einstein's equations. The corresponding gravitational entropy functional is

$$S_{\rm GR}[n^a] = \int_{\mathcal{V}} \frac{d^D x}{8\pi} \left(\nabla_a n^b \nabla_b n^a - (\nabla_c n^c)^2 \right) \tag{2.40}$$

Interestingly, the integrand in $S_{\rm GR}$ has the $Tr(K^2) - (TrK)^2$ structure. If we think of the $D=4$ spacetime being embedded in a sufficiently large k-dimensional *flat* spacetime we can obtain the same structure using the Gauss–Codazzi equations relating the (zero) curvature of k-dimensional space with the curvature of spacetime. As mentioned earlier, one can express any vector field n^a in terms of a set of basis vector fields n_A^a. Therefore, one can equivalently think of the functional $S_{\rm GR}$ as given by

$$S_{\rm GR}[n_A^a] = \int_{\mathcal{V}} \frac{d^D x}{8\pi} \left(\nabla_a n_I^b \nabla_b n_J^a - \nabla_c n_I^c \nabla_a n_J^a \right) P^{IJ} \tag{2.41}$$

where P^{IJ} is a suitable projection operator. It is not clear whether the embedding approach leads to any better understanding of the formalism; in particular, it does not seem to generalize in a natural fashion to Lanczos–Lovelock models.

The next-order term (which arises from the Gauss–Bonnet Lagrangian) is:

$$\overset{(2)}{P}{}^{ab}_{cd} = \frac{1}{16\pi} \frac{1}{2} \delta^{ab a_3 a_4}_{cd b_3 b_4} R^{b_3 b_4}_{a_3 a_4} = \frac{1}{8\pi} \left(R^{ab}_{cd} - G^a_c \delta^b_d + G^b_c \delta^a_d + R^a_d \delta^b_c - R^b_d \delta^a_c \right) \tag{2.42}$$

and similarly for all higher-order terms. None of them can contribute in $D=4$, so we get Einstein's theory as the unique choice if we assume $D=4$. If we assume that P^{abcd} is to be built *only* from the metric, then this choice is unique in all D.

Harold: You originally gave a motivational argument as to why this S should be thought of as entropy. As far as the variational principle is concerned, this identification does not seem to play a crucial role.

Me: It does rather indirectly. To see this, you only need to consider the form of S when the equations of motion are satisfied. First of all, Eq. (2.34) shows that when the equations of motion hold the total entropy of a bulk region is entirely on its boundary, which is nice. Further, if you evaluate this boundary term

$$-S|_{\rm on-shell} = 4 \int_{\partial\mathcal{V}} d^{D-1} x k_a \sqrt{h} \left(P^{abcd} n_c \nabla_b n_d \right) \tag{2.43}$$

(where we have manipulated a few indices using the symmetries of P^{abcd}) in the case of a *stationary* horizon which can be locally approximated as Rindler spacetime, one gets exactly the Wald entropy of the horizon [3]. This is one clear reason why we can think of S as entropy.

Harold: Is this entropy positive definite? Do you worry about that?
Me: I don't worry about that (yet!). In $D=4$, I can prove that the on-shell entropy is positive definite. But if one is dealing with a Lanczos–Lovelock model with horizons attributed to Wald entropy, it is known that [29] even on-shell entropy will not be positive definite for the whole range of parameters. Maybe this will put additional restrictions on the kind of gravitational theories which are physically reasonable. (This approach has uncovered several other issues related to entropy, quasi-normal modes, etc. and even a possibility of entropy being quantized [30], but all that will take us far afield.) At present these questions are open.
Harold: Usually in an action principle one varies all the degrees of freedom in any order one chooses. But in your extremum principle, we are expected to vary only n^a. How would I get equations of motion for matter in this approach?
Me: That is not a problem. At the classical level, the field equations are already contained in the condition $\nabla_a T^{ab}=0$ which I impose (with an intriguing interpretation, which you may not want to buy, that this is the constancy of elastic constants!). If you want to do quantum field theory in a curved spacetime, you can again use these field equations in the Heisenberg picture. The only question is when you insist that you need to do path integral quantizations of the matter fields. Then you have to, of course, vary n^a first and get the classical equations for gravity because the expression in Eq. (2.24) is designed as an entropy functional. But after you have done that and written down the field equations for gravity, you can do the usual variation of matter Lagrangian in a given curved spacetime and get the standard equations [3].

2.5 Comparison with the conventional perspective and further comments

Harold: I also notice that while the vector field n_a in LIF, to the lowest order, has no bulk dynamics, if you consider the integral over the Lagrangian in a small region in LIF, you will get a surface contribution. That seems strange, too.
Me: Not really. I am not surprised by S picking up just a surface contribution because – even in the conventional approach – Einstein–Hilbert action is holographic in a specific sense of the word [31] and does exactly that.
Harold: Maybe this is a good time to sort this out. You mentioned earlier that the democracy of observers and their right to do physics in spite of the existence of horizons has something to do with the holography of action. I have no idea what you are talking about here!
Me: Let me elaborate. We said that there should exist a mechanism which will encode the information in the region $\mathcal{V}$ which is inaccessible to a particular observer at the boundary $\partial\mathcal{V}$ of that region [14]. One possible way of ensuring this is to

add a suitable boundary term to the action principle which will provide additional information content for observers who perceive a horizon. Such a procedure leads to three immediate consequences.

First, if the theory is generally covariant, so that observers with horizons (like, for example, uniformly accelerated observers using a Rindler metric) need to be accommodated in the theory, such a theory *must* have an action functional that contains a surface term. The generally covariant action in Einstein's theory did contain a surface term. The present approach explains the logical necessity for such a surface term in a generally covariant theory which was not evident in the standard approach.

Harold: That's interesting. You are now claiming that there is a connection between the following three facts. (1) The theory for gravity is built from a generally covariant Lagrangian. (2) In a geometrical theory of gravity, horizons are inevitable but general covariance demands that all observers have an equal right to describe physics. (3) Observers whose information is blocked by a horizon should still be able to somehow get around this fact with the information encoded on the boundary. Therefore, the Lagrangian must have a boundary term. Viewed this way, it appears natural that the only generally covariant scalar Lagrangian proportional to R leads to a surface term in the action. But how does the surface term know what is going on in the bulk?

Me: That is the second point. If the surface term has to encode the information which is blocked by the horizon, then there must exist a simple relation between the bulk term and the surface term in the action and hence you cannot choose just any scalar. This is indeed the case for the Einstein–Hilbert action; there is a peculiar (unexplained) relationship between L_{bulk} and L_{sur}:

$$\sqrt{-g}L_{\text{sur}} = -\partial_a \left(g_{ij} \frac{\partial \sqrt{-g} L_{\text{bulk}}}{\partial(\partial_a g_{ij})} \right) \tag{2.44}$$

This shows that the Einstein–Hilbert gravitational action is 'holographic' with the same information being coded in both the bulk and surface terms.

In fact, in any local region around an event, it is the surface term which contributes to the action at the lowest order. In the neighborhood of any event, the Riemann normal coordinates in which $g \simeq \eta + R\,x^2$, $\Gamma \simeq R\,x$. In the gravitational Lagrangian $\sqrt{-g}R \equiv \sqrt{-g}L_{\text{bulk}} + \partial_a P^a$ with $L_{\text{bulk}} \simeq \Gamma^2$ and $\partial P \simeq \partial \Gamma$, the L_{bulk} term vanishes in this neighborhood while $\partial_a P_b \simeq R_{ab}$ leading to

$$\int_{\mathcal{V}} d^4x R\sqrt{-g} \approx \int_{\mathcal{V}} d^4x \partial_a P^a \approx \int_{\partial\mathcal{V}} d^3x n_a P^a \tag{2.45}$$

showing that in a small region around the event in the Riemann normal coordinates, gravitational action can be reduced to a pure surface term.

Harold: So in this perspective, you also expect the surface term to be related to the information content blocked by the horizon, right?
Me: Indeed. That is the third point. If the surface term encodes information which is blocked by the horizon, then it should actually lead to the entropy of the horizon. In other words, we should be able to compute the horizon entropy by evaluating the surface term. This is indeed true and can easily be demonstrated [31]. The surface term does give the horizon entropy for any metric for which near-horizon geometry has the Rindler form.

This explains another deep mystery in the conventional approach. In the usual approach, we *ignore* the surface term completely (or cancel it with a counter-term) and obtain the field equation from the bulk term in the action. Any solution to the field equation obtained by this procedure is logically independent of the nature of the surface term. But we find that when the *surface term* (which was ignored) is evaluated at the horizon that arises in any given solution, it does correctly give the entropy of the horizon! This is possible only because there is a relationship, given by Eq. (2.44), between the surface term and the bulk term which is again an unexplained feature in the conventional approach to gravitational dynamics. Since the surface term has the thermodynamic interpretation as the entropy of horizons, and is related holographically to the bulk term, we are again led to an indirect connection between spacetime dynamics and horizon thermodynamics.
Harold: I agree that these results are extremely mysterious in the conventional approach, now that you have brought it up. I have not seen Eq. (2.44) mentioned, let alone discussed in any work (other than yours, of course). I presume this is one of what you call the ‘algebraic accidents’. But if your ideas about Lanczos–Lovelock theory being the natural candidate in D dimensions are correct then the same ‘algebraic accident’ should occur in Lanczos–Lovelock theories as well, right?
Me: Yes. In fact it does – which is gratifying – and acts as a nontrivial consistency check on my alternative perspective. One can show that the surface and bulk terms of all Lanczos–Lovelock theories satisfy an equation similar to Eq. (2.44). Of course, since it wasn’t noticed for Hilbert action, nobody bothered about Lanczos–Lovelock action till we [5] unearthed it.

One can provide a simple, yet very general, proof of the connection between entropy and the surface term in action in any static spacetime. Such a spacetime will have a Killing vector ξ^a and a corresponding Noether current. Taking the J^0 component of Eq. (2.12) and writing $J^0 = \nabla_b J^{0b}$, we obtain

$$L = \frac{1}{\sqrt{-g}}\partial_\alpha\left(\sqrt{-g}\,J^{0\alpha}\right) - 2E^0_0 \tag{2.46}$$

Only spatial derivatives contribute in the first term on the right-hand side when the spacetime is static. This relation shows that the action obtained by integrating

$L\sqrt{-g}$ will generically have a surface term related to J^{ab}. (In Einstein gravity Eq. (2.46) will read $L = 2R^0_0 - 2G^0_0$; our result generalizes the fact that R^0_0 can be expressed as a total divergence in static spacetimes.) This again illustrates, in a very general manner, why the surface terms in the action functional lead to horizon entropy. In fact, Eq. (2.46) can be integrated to show that in any static spacetime with a bifurcation horizon, the action can be interpreted as the free energy which generalizes a result known in Einstein gravity to Lanczos–Lovelock models.

Harold: What are the other key algebraic accidents in the conventional approach which your perspective throws light on?

Me: There are several, but let me describe one that is really striking (and was first discussed in [32]). Consider a static, spherically symmetric horizon, in a spacetime described by a metric

$$ds^2 = -f(r)c^2dt^2 + f^{-1}(r)dr^2 + r^2 d\Omega^2 \tag{2.47}$$

Let the location of the horizon be given by the simple zero of the function $f(r)$, say at $r = a$. The Taylor series expansion of $f(r)$ near the horizon $f(r) \approx f'(a)(r-a)$ shows that the metric reduces to the Rindler metric near the horizon in the $r-t$ plane with the surface gravity $\kappa = (c^2/2)f'(a)$. Then, an analytic continuation to imaginary time allows us to identify the temperature associated with the horizon to be

$$k_B T = \frac{\hbar c f'(a)}{4\pi} \tag{2.48}$$

where we have introduced the normal units. The association of temperature in Eq. (2.48) with the metric in Eq. (2.47) only requires the conditions $f(a) = 0$ and $f'(a) \neq 0$. The discussion so far has not assumed anything about the dynamics of gravity or Einstein's field equations.

We shall now take the next step and write down the Einstein equation for the metric in Eq. (2.47), which is given by $(1-f) - rf'(r) = -(8\pi G/c^4)Pr^2$ where $P = T^r_r$ is the radial pressure. When evaluated on the horizon $r = a$ we get the result

$$\frac{c^4}{G}\left[\frac{1}{2}f'(a)a - \frac{1}{2}\right] = 4\pi P a^2 \tag{2.49}$$

If we now consider two solutions to Einstein's equations differing infinitesimally in the parameters such that horizons occur at two different radii a and $a + da$, then multiplying Eq. (2.49) by da, we get

$$\frac{c^4}{2G}f'(a)a\,da - \frac{c^4}{2G}da = P(4\pi a^2 da) \tag{2.50}$$

The right-hand side is just PdV, where $V=(4\pi/3)a^3$ is what is called the areal volume which is the relevant quantity when we consider the action of pressure on a surface area. In the first term, we note that $f'(a)$ is proportional to the horizon temperature in Eq. (2.48). Rearranging this term slightly and introducing a $\hbar$ factor *by hand* into an otherwise classical equation to bring in the horizon temperature, we can rewrite Eq. (2.50) as

$$\underbrace{\frac{\hbar c f'(a)}{4\pi}}_{k_BT}\underbrace{\frac{c^3}{G\hbar}d\left(\frac{1}{4}4\pi a^2\right)}_{dS}\underbrace{-\frac{1}{2}\frac{c^4 da}{G}}_{-dE}=\underbrace{Pd\left(\frac{4\pi}{3}a^3\right)}_{PdV} \tag{2.51}$$

The labels below the equation indicate a natural – and unique – interpretation for each of the terms and the whole equation now becomes $TdS=dE+PdV$, allowing us to read off the expressions for entropy and energy:

$$S=\frac{1}{4L_P^2}(4\pi a^2)=\frac{1}{4}\frac{A_H}{L_P^2};\quad E=\frac{c^4}{2G}a=\frac{c^4}{G}\left(\frac{A_H}{16\pi}\right)^{1/2} \tag{2.52}$$

where A_H is the horizon area and $L_P^2=G\hbar/c^3$. The result shows that Einstein's equations can be reinterpreted as a thermodynamic identity for a virtual displacement of the horizon by an amount da.

Harold: I suppose the uniqueness of the factor $P(4\pi a^2)da$, where $4\pi a^2$ is the proper area of a surface of radius a in spherically symmetric spacetimes, implies that we cannot carry out the same exercise by multiplying Eq. (2.49) by some other arbitrary factor $F(a)da$ instead of just da in a natural fashion. This, in turn, uniquely fixes both dE and the combination TdS. The product TdS is classical and is independent of $\hbar$, and hence, we can determine T and S only within a multiplicative factor. The only place you introduced $\hbar$ by hand is in using the Euclidean extension of the metric to fix the form of T and thus S. Right?

Me: Yes. With that I can remove the ambiguity in the overall multiplicative factor. So, given the structure of the metric in Eq. (2.47) and Einstein's equations, we can determine T, S and E uniquely. The fact that $T\propto\hbar$ and $S\propto 1/\hbar$ is analogous to the situation in classical thermodynamics in contrast with statistical mechanics. The TdS in thermodynamics is independent of Boltzmann's constant, while statistical mechanics will lead to $S\propto k_B$ and $T\propto 1/k_B$.

Harold: That is a bit mind-boggling. Usually, the rigorous way of obtaining the temperature of a horizon – say, a black hole horizon – is by studying a quantum field in the externally specified metric. *You never need to specify whether the metric is a solution to Einstein's equations.* Now you are telling me that the same result arises *without any reference to an externally specified quantum field theory but on* *using Einstein's equations on the horizons*. How do the Einstein equations know that there is a temperature, entropy, etc.?

Me: Yes. That is the algebraic coincidence. More sharply stated, we have no explanation as to why an equation like Eq. (2.51) should hold in classical gravity, if we take the conventional route. This strongly suggests that the association of entropy and temperature with a horizon is quite fundamental and is actually connected with the dynamics (encoded in Einstein's equations) of the gravitational field. The fact that quantum field theory in a spacetime with horizon exhibits thermal behaviour should then be thought of as a *consequence* of a more fundamental principle.

Harold: If so, the idea should also have a more general validity. Does it?

Me: Yes. One can again show that the field equations of more general theories of gravity (like in Lanczos–Lovelock models) also reduce to the same thermodynamic identity $TdS = dE + PdV$ when evaluated on the horizon. This has now been demonstrated [33] for an impressively wide class of models like (i) the stationary axisymmetric horizons and (ii) evolving spherically symmetric horizons in Einstein gravity, (iii) static spherically symmetric horizons and (iv) dynamical apparent horizons in Lovelock gravity, (v) three-dimensional BTZ black hole horizons, (vi) FRW cosmological models in various gravity theories and (vii) even [34] in the case of Horava–Lifshitz gravity. It is not possible to understand, in the conventional approach, why the field equations should encode information about horizon thermodynamics.

Harold: This is a fairly strong argument in favour of a thermodynamic underpinning for the dynamics of gravity. But before I accept that *in toto*, I need to convince myself that there is no simpler explanation for this result. I accept that none is given in the literature, but how about the standard first law of black hole thermodynamics? Your relation looks similar to it so I wonder whether there is a connection.

Me: No. We are talking about very different things. In general, in spite of the superficial similarity, Eq. (2.51) is *different* from the conventional first law of black hole thermodynamics due to the presence of the PdV term. The difference is easily seen, for example, in the case of a Reissner–Nordstrom black hole for which $T^r_r = P$ is nonzero due to the presence of a nonzero electromagnetic energy–momentum tensor in the right-hand side of Einstein's equations. If a *chargeless* particle of mass dM is dropped into a Reissner–Nordstrom black hole, then the standard first law of black hole thermodynamics will give $TdS = dM$. But in Eq. (2.51), the energy term, defined as $E \equiv a/2$, changes by $dE = (da/2) = (1/2)[a/(a-M)]dM \neq dM$. It is easy to see, however, that for the Reissner–Nordstrom black hole, the combination $dE + PdV$ is precisely equal to dM, making sure $TdS = dM$. So we need the PdV term to get $TdS = dM$ from Eq. (2.51) when a *chargeless* particle is dropped into a Reissner–Nordstrom black hole. More generally, if da arises due to changes dM and dQ, it is easy to show that Eq. (2.51) gives $TdS = dM - (Q/a)dQ$ where the

second term arises from the electrostatic contribution. This ensures that Eq. (2.51) is perfectly consistent with the standard first law of black hole dynamics in those contexts in which both are applicable but $dE \neq dM$ in general. You would also have realized that the way Eq. (2.51) was derived is completely local and quite different from the way one obtains the first law of black hole thermodynamics.
Harold: Yes, I see that. It appears that gravitational field equations and their solutions with horizons have a deeper connection with thermodynamics than is apparent. In fact, I believe you would claim the thermodynamic perspective is more fundamental than the field equations describing gravity.
Me: Precisely. That is why I spent a lot of time explaining the thermodynamic motivation in Sections 2.2 and 2.3 while I could have derived the field equations just by extremizing the expression in Eq. (2.24). But in a way everything else is just motivational if you are willing to accept the perspective based on Eq. (2.24) as fundamental.
Harold: That brings up the question you promised a discussion on. What about the cosmological constant [35]? In the conventional approach, one introduces it as a term in the gravitational Lagrangian. You don't have any such term in Eq. (2.24) but nevertheless the cosmological constant appears in your final equations!
Me: Yes, and I would claim that this is another very attractive feature of this new perspective. In the standard approach, one starts with an action

$$\mathcal{A}_{\rm tot} = \int d^D x \sqrt{-g}\, (L_{\rm grav} + L_m) \tag{2.53}$$

and varies (i) the matter degrees of freedom to obtain the equations of motion for matter and (ii) the metric g^{ab} to obtain the field equations of gravity. The equations of motion for matter remain invariant if one adds a constant, say, $-\rho_0$ to the matter Lagrangian, which is equivalent to adding a constant ρ_0 to the Hamiltonian density of the matter sector. Physically, this symmetry reflects the fact that the zero level of the energy is arbitrary in the matter sector and can be set to any value without leading to observable consequences. However, gravity breaks this symmetry which the matter sector has. A shift $L_m \to L_m - \rho_0$ will change the energy–momentum tensor T^a_b which acts as the source of gravity by a term proportional to $\rho_0 \delta^a_b$. Therefore, having a nonzero baseline for energy density of matter is equivalent to a theory with cosmological constant which – in turn – will lead to observable consequences. If we interpret the evidence for dark energy in the Universe (see [36]; for a critical look at the data, see [37] and the references cited therein) as due to the cosmological constant, then its value has to be fine-tuned to enormous[4]

[4] This is, of course, the party line. But it might help to get some perspective on how enormous, the 'enormous' really is. To begin with note that, the sensible particle physics convention considers ratios of length (or energy)

accuracy to satisfy the observational constraints. It is not clear why a particular parameter in the low-energy matter sector has to be fine-tuned in such a manner.

In the alternative perspective described here, the functional in Eq. (2.24) is clearly invariant under the shift $L_m \to L_m - \rho_0$ or equivalently, $T_{ab} \to T_{ab} + \rho_0 g_{ab}$, since it only introduces a term $-\rho_0 n_a n^a = 0$ for any null vector n_a. In other words, one *cannot* introduce the cosmological constant as a low-energy parameter in the action in this approach. We saw, however, that the cosmological constant reappears as *an integration constant* when the equations are solved. The integration constants which appear in a particular solution have a completely different conceptual status compared to the parameters which appear in the action describing the theory. It is much less troublesome to choose a fine-tuned value for a particular integration constant in the theory if observations require us to do so. From this point of view, the cosmological constant problem is considerably less severe when we view gravity from the alternative perspective.

Harold: I suppose you succeed in having the extra symmetry under the shift $T_{ab} \to T_{ab} + \rho_0 g_{ab}$ because you are not treating the metric as a dynamical variable, right?

Me: Right. In fact, one can state a stronger result [38]. Consider any model of gravity satisfying the following three conditions. (1) The metric is varied in a local action to obtain the equations of motion. (2) We demand full general covariance of the equations of motion. (3) The equations of motion for the matter sector are invariant under the addition of a constant to the matter Lagrangian. Then, we can prove a 'no-go' theorem that the cosmological constant problem cannot be solved in such a model [38]. The proof is elementary. Our demand (2) for general covariance requires the matter action to be an integral over $\mathcal{L}_{\rm matter}\sqrt{-g}$. Demand (3) now allows us to add a constant Λ, say, to $\mathcal{L}_{\rm matter}$, leading to a coupling $\Lambda\sqrt{-g}$ between Λ and the metric g_{ab}. By our demand (1), when we vary g_{ab} the theory will couple to Λ through a term proportional to Λg_{ab}, thereby introducing an arbitrary cosmological constant into the theory.

The power of the above 'no-go theorem' lies in its simplicity! It clearly shows that we cannot solve the cosmological constant problem unless we drop one of the three demands listed in the above paragraph. Of these, we do not want to sacrifice the general covariance encoded in (2); neither do we have a handle on the low-energy matter Lagrangian so we cannot avoid (3). So the only hope we have is to introduce an approach in which gravitational field equations are obtained from varying some

scales and not their *squares* as cosmologists are fond of doing. This leads to $(L_P/L_\Lambda) \sim 10^{-61}$ instead of the usual $\Lambda L_P^2 \sim 10^{-122}$. In the standard model of particle physics the ratio between Planck scale and neutrino mass scale is $10^{19}\,\mathrm{GeV}/10^{-2}\,\mathrm{eV} \sim 10^{30}$, for which we have no theoretical explanation. So when we worry about the fine-tuning of the cosmological constant without expressing similar worries about the standard model of particle physics, we are essentially assuming that 10^{30} is not a matter for concern but 10^{61} is. This subjective view is defensible but needs to be clearly understood.

degrees of freedom other than g_{ab} in a maximization principle. This suggests that the so-called cosmological constant problem has its roots in our misunderstanding of the nature of gravity.

Harold: I thought that any spin-2 long-range field h_{ab} (arising, for example, in the linear perturbation around flat spacetime through $g_{ab} = \eta_{ab} + h_{ab}$) obeying the principle of equivalence has to couple to T_{ab} generically through a term in the action $T^{ab}h_{ab}$. But in your model, this does not seem to happen.

Me: That's correct. It is sometimes claimed that a spin-2 graviton in the linear limit *has to* couple to T_{ab} in a universal manner, in which case, one will have the graviton coupling to the cosmological constant. In our approach, the linearized field equations for the spin-2 graviton field $h_{ab} = g_{ab} - \eta_{ab}$, in a suitable gauge, will be $(\Box h_{ab} - T_{ab})n^a n^b = 0$ for all null vectors n^a. This equation is still invariant under $T_{ab} \to T_{ab} + \rho_0 g_{ab}$, showing that the graviton does *not* couple to the cosmological constant.

Harold: But can you predict the observed value of the cosmological constant in your approach?

Me: Alas, no. But I claim providing a mechanism in which the *bulk cosmological constant decouples from gravity* is a major step forward. If the cosmological constant was strictly zero, my perspective has a natural explanation for it – which no one else had! It was always thought that this should arise from some unknown symmetry and I have provided you with a model which has such symmetry. I believe the small value of the observed cosmological constant arises from nonperturbative quantum gravitational effects at the next order, but I don't have a fully satisfactory model. (See, however, [39].)

2.6 Summary and outlook

Harold: We have covered a lot of ground, some of which is purely technical while the rest is conceptual or interpretational. In your mind the distinction may be unimportant but others will react differently to results which can be rigorously proved compared to interpretational aspects, however elegant the latter may be. Maybe you would care to separate them out and provide a summary?

Me: Fine. From a purely algebraic point of view, without bringing in any physical interpretation or motivation, we can prove the following mathematical results.

- Consider a functional of null vector fields $n^a(x)$ in an arbitrary spacetime given by Eq. (2.24) [or, more generally, by Eq. (2.36)]. Demanding that this functional is an extremum for all null vectors n^a leads to the field equations for the background geometry given by $(2E_{ab} - T_{ab})n^a n^b = 0$ where E_{ab} is given by Eq. (2.32) [or, more generally, by Eq. (2.2)]. Thus, field equations in a wide class of theories of gravity can be obtained from an extremum principle without varying the metric as a dynamical variable.

- These field equations are invariant under the transformation $T_{ab} \to T_{ab} + \rho_0 g_{ab}$, which relates to the freedom of introducing a cosmological constant as an integration constant in the theory. Further, this symmetry forbids the inclusion of a cosmological constant term in the variational principle by hand as a low-energy parameter. *That is, we have found a symmetry which makes the bulk cosmological constant decouple from gravity.* When linearized around flat spacetime, the graviton inherits this symmetry and does not couple to the cosmological constant.
- On-shell, the functional in Eq. (2.24) [or, more generally, Eq. (2.36)] contributes only on the boundary of the region. When the boundary is a horizon, this terms gives precisely the Wald entropy of the theory.

Harold: It is remarkable that you can derive not only Einstein's theory uniquely in $D = 4$ but even Lanczos–Lovelock theory in $D > 4$ from an extremum principle involving the null normals *without varying g_{ab} in an action functional!* I also see from Eq. (2.33) that in the case of Einstein's theory, you have a Lagrangian $n^a(\nabla_{[a}\nabla_{b]})n^b$ for a vector field n^a which becomes vacuous in flat spacetime in which covariant derivatives become partial derivatives. There is clearly no dynamics in n^a but they do play a crucial role. So I see that you can get away without ever telling me what the null vectors n^a actually mean because they disappear from the scene after serving their purpose. While this may be mathematically clever, it is very unsatisfactory physically. You may not have a rigorous model for these degrees of freedom, but what is your picture?
Me: My picture is made of the following ingredients, each of which seems reasonable but far from having rigorous mathematical justification at this stage.

- Assume that the spacetime is endowed with certain microscopic degrees of freedom capable of exhibiting thermal phenomena. This is just the Boltzmann paradigm: *if one can heat it, it must have microstructure!*; and one can heat up a spacetime.
- Whenever a class of observers perceive a horizon, they are 'heating up the spacetime' and the degrees of freedom close to a horizon participate in a very *observer-dependent* thermodynamics. Matter which flows close to the horizon (say, within a few Planck lengths of the horizon) transfers energy to these microscopic, near-horizon, degrees of freedom *as far as the observer who sees the horizon is concerned*. Just as the entropy of a normal system at temperature T will change by $\delta E/T$ when we transfer to it an energy δE, here also an entropy change will occur. (A freely falling observer in the same neighbourhood, of course, will deny all these!)

- We proved that when the field equations of gravity hold, one can interpret this entropy change in a purely geometrical manner involving the Noether current. From this point of view, the normals n^a to local patches of null surfaces are related to the (unknown) degrees of freedom that can participate in the thermal phenomena involving the horizon.
- Just as demanding the validity of special relativistic laws with respect to all freely falling observers leads to the kinematics of gravity, demanding the local entropy balance in terms of the thermodynamic variables as perceived by local Rindler observers leads to the field equations of gravity in the form $(2E_{ab} - T_{ab})n^a n^b = 0$.

Harold: It is an interesting picture, but is *totally observer-dependent* right? A local Rindler observer or an observer outside a black hole horizon might attribute all kinds of thermodynamics and entropy changes to the horizons she perceives. But an inertial observer or an observer falling through the Schwarzschild horizon will see none of these phenomena.
Me: Exactly. I claim we need to accept the fact that a whole lot of thermodynamic phenomena need now to be thought of as observer-dependent. For example, if you throw some hot matter on to a Schwarzschild black hole, then when it gets to a few Planck lengths away from the horizon and hovers around it, I expect it to interact with the microscopic horizon degrees of freedom *as far as an outside observer is concerned*. After all, such an observer would claim that all matter stays arbitrarily close but outside the horizon for all eternity. A freely falling observer through the horizon will have a completely different picture, but we have learnt to live with this dichotomy as far as elementary kinematics goes. I think we need to do the same as regards thermodynamics and quantum processes.
Harold: In a way, every key progress in physics has involved realizing that something we thought of as absolute is not absolute. With special relativity it was the flow of time and with general relativity it was the concept of global inertial frames and when we brought in quantum fields in curved spacetime it was the notion of particles.
Me: And the notion of temperature, don't forget that. We now know that the temperature attributed to even the vacuum state depends on the observer. We need to go further and integrate the entire thermodynamic machinery – involving highly excited semi-classical states, say, cups of tea with (what we believe to be) 'real' temperature – with this notion of LRFs having their own temperature. I don't think this has been done in a satisfactory manner yet [21].
Harold: So, what next? What is the 'to do list'?
Me: To name a few that come to mind, I can list them as technical ones and conceptual ones. On the technical side:

(i) It would be nice to make the notion of LRF and the horizon a bit more rigorous. For example, the idea of an approximate Killing vector could be made more precise and one might like to establish the connection between locality, which is apparent in the Euclidean sector and causality, which is apparent in the Lorentzian section that has light cones. By and large, one would like to make rigorous the use of LRFs by, say, computing the next-order corrections.

(ii) A lot more can be done to clarify the observer dependence of the entropy. (The study of horizon thermodynamics makes one realize that one does not really quite understand what entropy is!) It essentially involves *exact* computation of $(S_1 - S_0)$, where S_1 and S_0 are the entropies attributed to the excited and ground state by a Rindler observer. This should throw more light on the expression for δS used in Eq. (2.4). In particular, it would be nice to have a detailed model which shows why δS involves the combination of δE of matter and T of the horizon. One would then use these insights to understand why Eq. (2.21) actually represents the relevant entropy functional for matter for arbitrary T_{ab}.

(iii) It will be nice to have a handle on the positivity or otherwise of the entropy functional used in Eq. (2.24).

These technical issues, I believe, can be tackled in a more or less straightforward manner, though the mathematics can be fairly involved. But as I said, they are probably not crucial to the alternative perspective or its further progress. The latter will depend on more serious conceptual issues, some of which are the following:

(i) How come the microstructure of spacetime exhibits itself indirectly through the horizon temperature even at scales much larger than Planck length? I believe this is because the event horizon works as some kind of magnifying glass allowing us to probe trans-Planckian physics [40], but this notion needs to be made more precise.

(ii) How does one obtain the expression for entropy in Eq. (2.24) from some microscopic model? In particular, such an analysis – even with a toy model – should throw more light on why normals to local patches of null surfaces play such a crucial role as effective degrees of freedom in the long-wavelength limit. Of course, such a model should also determine the expression for P^{abcd} and get the metric tensor and spacetime as derived concepts – a fairly tall order! (This is somewhat like obtaining the theory of elasticity starting from a microscopic model for a solid, which, incidentally, is not a simple task either.)

Harold: But what about the 'deep questions' like, for example, the physics near the singularities? Since you get the same field equations as anybody else does, you will have the same solutions, same singularities, etc.

Me: I told you that I am not doing statistical mechanics (which would be the full quantum theory) of spacetime but only thermodynamics. To answer issues related to singularities, etc., one actually needs to discover the statistical mechanics underlying the thermodynamic description I have presented here. We can have another chat, after I figure out the statistical mechanics of the spacetime microstructure!

Acknowledgements

The questions of Harold mostly represent issues raised by several colleagues – far too numerous to name individually – in my lectures, discussions, etc. I thank all of them for helping me to sharpen the ideas. I also thank A. D. Patel, K. Subramanian, Sudipta Sarkar, Aseem Paranjape, D. Kothawala and Sunu Engineer for several rounds of discussions over the past many years. I would also like to thank the participants of the 'Foundations of Space and Time' meeting at Cape Town, 10–14 August 2009, for discussions and comments. Finally, it is a pleasure to thank G. F. R. Ellis, A. Weltman and Jeff Murugan for organizing such a stimulating conference and providing great hospitality.

References

[1] J. Schwinger, *Particles, Sources and Fields*, Volume I (Perseus Books, 1998).
[2] T. Padmanabhan, Entropy density of spacetime and thermodynamic interpretation of field equations of gravity in any diffeomorphism invariant theory [arXiv:0903.1254]; T. Padmanabhan, Entropy density of spacetime and gravity: a conceptual synthesis, to appear in *IJMPD* (2009).
[3] T. Padmanabhan, *Gen. Rel. Grav.*, **40**, 529–64 (2008) [arXiv:0705.2533]; T. Padmanabhan, A. Paranjape, *Phys. Rev.* **D75**, 064004 (2007).
[4] T. Padmanabhan, *Adv. Sci. Lett.*, **2**, 174 (2009) [arXiv:0807.2356]; T. Padmanabhan, Gravity – the inside story [First Prize Essay, Gravity Research Foundation Essay Contest, 2008], *Gen. Rel. Grav.*, **40**, 2031 (2008).
[5] A. Mukhopadhyay, T. Padmanabhan, *Phys. Rev.*, **D74**, 124023 (2006) [hep-th/0608120].
[6] N. Deruelle, J. Katz, and S. Ogushi, *Class. Quant. Grav.*, **21**, 1971 (2004), [gr-qc/0310098]; G. L. Cardoso, B. de Wit, T. Mohaupt, arXiv:hep-th/9904005v2; T. Padmanabhan, *Gravitation: Foundations and Frontiers* (Cambridge University Press, 2009).
[7] For a sample of other approaches, see e.g., G. E. Volovik, *The Universe in a Helium Droplet* (Oxford University Press, 2003); T. Jacobson, *Phys. Rev. Lett.*, **75**, 1260 (1995) [gr-qc/9504004]; L. Sindoni *et al.*, arXiv:0909.5391; B. L. Hu, arXiv:0903.0878; M. Visser, *Mod. Phys. Lett.*, **A17**, 977 (2002) [gr-qc/0204062]; C. Barcelo *et al.*, *Int. J. Mod. Phys.* **D10** (2001) 799 [gr-qc/0106002]; C. G. Huang, J. R. Sun, gr-qc/0701078; J. Makela, gr-qc/0701128.
[8] A. D. Sakharov, *Sov. Phys. Dokl.*, **12**, 1040 (1968).
[9] T. Padmanabhan, *Dark Energy: Mystery of the Millennium*, Albert Einstein Century International Conference, Paris, 18–22 July 2005, AIP Conference Proceedings 861, pp. 858–66 [astro-ph/0603114].
[10] P. C. W. Davies, *J. Phys. A*, **8**, 609 (1975); W. G. Unruh, *Phys. Rev.*, **D14**, 870 (1976).
[11] S. W. Hawkings, *Commun. Math. Phys.*, **43**, 199–220 (1975).

[12] T. Padmanabhan, *Phys. Rep.*, **406**, 49 (2005) [gr-qc/0311036]; *AIP Conference Proceedings*, **989**, 114 (2007) [arXiv:0706.1654].

[13] C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation* (W.H. Freeman, 1973), chapter 7.

[14] T. Padmanabhan, A. Patel, *Role of Horizons in Semiclassical Gravity: Entropy and the Area Spectrum* [gr-qc/0309053]; T. Padmanabhan, *Gen. Rel. Grav.*, **35**, 2097–103 (2003) [Fifth Prize Essay; Gravity Research Foundation Essay Contest, 2003]; *Mod. Phys. Lett.* **A17**, 923 (2002) [gr-qc/0202078].

[15] See e.g., T. Roy Choudhury, T. Padmanabhan, *Gen. Rel. Grav.*, **39**, 1789 (2007) [gr-qc/0404091].

[16] D. Kothawala and T. Padmanabhan, Response of Unruh–De Witt detector with time-dependent acceleration (2009) [arXiv:0911.1017].

[17] T. Padmanabhan, *Mod. Phys. Lett*. A, **18**, 2903 (2003) [hep-th/0302068]; *Mod. Phys. Lett*. A **19**, 2637–43 (2004) [gr-qc/0405072].

[18] This is related to the famous question first posed by Wheeler to Bekenstein: What happens if you mix cold and hot tea and pour it down a horizon, erasing all traces of ‘crime’ in increasing the entropy of the world? This is based on what Wheeler told me in 1985, from *his* recollection of events; it is also mentioned in his book, J.A. Wheeler, *A Journey into Gravity and Spacetime* (Scientific American Library, 1990), p. 221. I have heard somewhat different versions from other sources.

[19] J.D. Bekenstein, *Phys. Rev.*, **D7**, 2333 (1973).

[20] H.S. Snyder, *Phys. Rev.*, **71**, 38 (1947); B.S. DeWitt, *Phys. Rev. Lett.*, **13**, 114 (1964); T. Yoneya, *Prog. Theor. Phys.*, **56**, 1310 (1976); T. Padmanabhan, *Ann. Phys.* (*N.Y.*), **165**, 38 (1985); *Class. Quant. Grav.*, **4**, L107 (1987); T. Padmanabhan, *Phys. Rev. Lett.*, **78**, 1854 (1997) [hep-th/9608182]; *Phys. Rev.*, **D57**, 6206 (1998); K. Srinivasan *et al.*, *Phys. Rev.*, **D58**, 044009 (1998) [gr-qc/9710104]; X. Calmet *et al.*, *Phys. Rev. Lett.*, **93**, 211101 (2004) [hep-th/0505144]; M. Fontanini *et al.*, *Phys. Lett. B*, **633**, 627 (2006) [hep-th/0509090]. For a review, see L. J. Garay, *Int. J. Mod. Phys.*, **A10**, 145 (1995).

[21] T. Padmanabhan, unpublished work in progress.

[22] D. Marolf, D. Minic, S. Ross, *Phys. Rev.*, **D69**, 064006 (2004).

[23] S. Kolekar, D. Kothawala, T. Padmanabhan, work in progress (2009).

[24] R.M. Wald, *Phys. Rev. D*, **48**, 3427 (1993) [gr-qc/9307038]; V. Iyer, R.M. Wald, *Phys. Rev. D*, **52**, 4430 (1995) [gr-qc/9503052].

[25] T. Padmanabhan, *Int. J. Mod. Phys.*, **D13**, 2293–8 (2004) [gr-qc/0408051]; *Int. J. Mod. Phys.*, **D14**, 2263–70 (2005) [gr-qc/0510015].

[26] L.D. Landau and E.M. Lifshitz, *Theory of Elasticity* (Pergamon Press, 2nd edition, 1981).

[27] S.-F. Wu *et al.* (2009), arXiv:0909.1367v2.

[28] C. Lanczos, *Z. Phys.*, **73**, 147 (1932); *Annals Math.*, **39**, 842 (1938); D. Lovelock, *J. Math. Phys.*, **12**, 498 (1971).

[29] M. Cvetic *et al.*, *Nucl. Phys.*, **B628**, 295 (2002) [hep-th/0112045]; S. Nojiri, S. D. Odintsov, *Phys. Rev.*, **D66**, 044012 (2002) [hep-th/0204112]; T. Clunan *et al.*, *Class. Quant. Grav.*, **21**, 3447 (2004) [gr-qc/0402044].

[30] D. Kothawala, T. Padmanabhan, S. Sarkar, *Phys. Rev.*, **D78**, 104018 (2008) [arXiv:0807.1481]; T. Padmanabhan, *Class. Quant. Grav.*, **21**, L1 (2004) [gr-qc/0310027]; *Class. Quan. Grav.*, **21**, 4485 (2004) [gr-qc/0308070]; T. Roy Choudhury, T. Padmanabhan, *Phys. Rev.*, **D69**, 064033 (2004) [gr-qc/0311064].

[31] T. Padmanabhan, *Gen. Rel. Grav.*, **34**, 2029–35 (2002) [gr-qc/0205090] [Second Prize Essay; Gravity Research Foundation Essay Contest, 2002] [gr-qc/0209088]; *Mod.*

Phys. Lett. **A17**, 1147 (2002) [hep-th/0205278]; *Braz. J. Phys.* (Special Issue), **35**, 362 (2005) [gr-qc/0412068]; T. Padmanabhan, Gravity: A New Holographic Perspective (Lecture at the International Conference on Einstein's Legacy in the New Millennium, December 2005), *Int. J. Mod. Phys.*, **D15**, 1659–75 (2006) [gr-qc/0606061].

[32] T Padmanabhan, *Class. Quant. Grav.*, **19**, 5387 (2002) [gr-qc/0204019].

[33] For a small sample, see e.g., D. Kothawala *et al.*, *Phys. Lett.*, **B652**, 338 (2007) [arXiv:gr-qc/0701002]; A. Paranjape *et al.*, *Phys. Rev.*, **D74**, 104015 (2006) [arXiv:hep-th/0607240]; D. Kothawala, T. Padmanabhan, *Phys. Rev.*, **D79**, 104020 (2009) [arXiv:0904.0215]; R. G. Cai *et al.*, *Phys. Rev.*, **D78**, 124012 (2008) [arXiv:0810.2610]; *Phys. Rev.*, **D75**, 084003 (2007) [arXiv:hep-th/0609128]; M. Akbar, R. G. Cai, *Phys. Lett.*, **B635**, 7 (2006) [arXiv:hep-th/0602156]; *Phys. Lett.*, **B648**, 243 (2007) [arXiv:gr-qc/0612089]; Y. Gong, A. Wang, *Phys. Rev. Lett.* **99**, 211301 (2007) [arXiv:0704.0793 [hep-th]]; S. F. Wu, G. H. Yang, P. M. Zhang [arXiv:0710.5394]; S. F. Wu, B. Wang, G. H. Yang, *Nucl. Phys. B*, **799**, 330 (2008) [arXiv:0711.1209]; S. F. Wu, B. Wang, G. H. Yang, P. M. Zhang, arXiv:0801.2688; J. Zhou *et al.*, arXiv:0705.1264; R.-G. Cai, L.-M. Cao, gr-qc/0611071; M. Akbar, hep-th/0702029; X.-H. Ge, hep-th/0703253; A. Sheykhi *et al.*, hep-th/0701198; G. Allemandi *et al.*, gr-qc/0308019.

[34] R. G. Cai, N. Ohta, *Horizon thermodynamics and gravitational field equations in Horava–Lifshitz gravity* [arXiv:0910.2307].

[35] For a review, see e.g., T. Padmanabhan, *Phys. Rep.* **380**, 235–320 (2003) [hep-th/0212290].

[36] S. J. Perlmutter *et al., Astrophys. J.*, **517**, 565 (1999); A. G. Reiss *et al., Astron. J.*, **116**, 1009 (1998); J. L. Tonry *et al., ApJ*, **594**, 1 (2003); B. J. Barris, *Astrophys. J.*, **602**, 571 (2004); A. G. Reiss *et al., Astrophys. J.*, **607**, 665 (2004).

[37] H. K. Jassal *et al.*, *Phys. Rev.* **D72**, 103503 (2005) [astro-ph/0506748]; [astro-ph/0601389]; T. Padmanabhan, T. Roy Choudhury, *MNRAS*, **344**, 823 (2003) [astro-ph/0212573]; T. Roy Choudhury, T. Padmanabhan, *Astron. Astrophys.*, **429**, 807 (2005), [astro-ph/0311622]; S. Nesseris, L. Perivolaropoulos, *JCAP*, **0702**, 025 (2007); Y. Wang, P. Mukherjee, *Phys. Rev.,* **D76**, 103533 (2007).

[38] T. Padmanabhan, Gravity's Immunity from Vacuum: The Holographic Structure of Semiclassical Action [Third Prize Essay; Gravity Research Foundation Essay Contest, 2006], *Gen. Rel. Grav.*, **38**, 1547–52 (2006); T. Padmanabhan, *Curr. Sci.*, **88**, 1057 (2005) [astro-ph/0411044]; *Int. J. Mod. Phys.*, **D15**, 2029 (2006) [gr-qc/0609012].

[39] T. Padmanabhan, *Class. Quan. Grav.*, **22**, L107–10 (2005) [hep-th/0406060]. For earlier attempts in a similar spirit, see T. Padmanabhan, *Class. Quant. Grav.*, **19**, L167 (2002) [gr-qc/0204020]; D. Sorkin, *Int. J. Theor. Phys.*, **36**, 2759 (1997); for related work, see G. E. Volovik, gr-qc/0405012; J. V. Lindesay *et al.*, astro-ph/0412477; Y. S. Myung, hep-th/0412224; J. D. Barrow, gr-qc/0612128; E. Elizalde *et al.*, hep-th/0502082.

[40] See e.g., T. Padmanabhan, *Phys. Rev. Lett.*, **81**, 4297 (1998) [hep-th/9801015]; *Phys. Rev.*, **D59**, 124012 (1999) [hep-th/9801138] and references cited therein.

3

Effective theories and modifications of gravity

C. P. BURGESS

We live at a time of contradictory messages about how successfully we understand gravity. General relativity seems to work very well in the Earth's immediate neighborhood, but arguments abound that it needs modification at very small and/or very large distances. This chapter tries to put this discussion into the broader context of similar situations in other areas of physics, and summarizes some of the lessons which our good understanding of gravity in the solar system has for proponents for its modification over very long and very short distances. The main message is that effective theories, in the technical sense of "effective," provide the natural language for testing proposals, and so are also effective in the colloquial sense.

3.1 Introduction

Einstein's recognition early last century that gravity can be interpreted as the curvature of space and time represented an enormous step forward in the way we think about fundamental physics. Besides its obvious impact for understanding gravity over astrophysical distances – complete with resolutions of earlier puzzles (like the detailed properties of Mercury's orbit) and novel predictions for new phenomena (like the bending of light and the slowing of clocks by gravitational fields) – its implications for other branches of physics have been equally profound.

These implications include many ideas we nowadays take for granted. One such is the universal association of fundamental degrees of freedom with fields (first identified for electromagnetism, but then cemented with its extension to gravity, together with the universal relativistic rejection of action at a distance). Another is

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

the recognition of the power of symmetries in the framing of physical law, and the ubiquity in particular of gauge symmetries in their description (again reinforcing the earlier discovery in electromagnetism). A third is the systematization of the belief that the physical content of Nature's laws should be independent of the variables used in their description, and the consequent widespread penetration of geometrical methods throughout physics.

But the study of general relativity (GR) and other interactions (like electromagnetism, and its later-discovered relatives: the weak and strong forces) have since drifted apart. Like ex-lovers who remain friends, for most of the last century practitioners in either area have known little of the nitty gritty of each other's day-to-day struggles, even as they read approvingly of their occasional triumphs in the popular press.

Over the years the study of both gravity and the other interactions has matured into precision science, with many impressive theoretical developments and observational tests. For gravity this includes remarkably accurate accounts of motion within the solar system, to the point that GR – through its use within the global positioning system (GPS) – is now an indispensable tool for engineers [35]. For the other interactions the successes include the development and testing of the Standard Model (SM), a unified framework for all known non-gravitational physics, building on the earlier successes of Quantum Electrodynamics (QED).

There is nevertheless a mounting chorus of calls for modifying general relativity, both at very short and very long distances. These arise due to perceived failures of the theory when applied over distances much different from those over which it is well tested. The failures at short distances are conceptual, to do with combining gravity with quantum effects. Those at long distances are instead observational, and usually arise as ways to avoid the necessity for introducing the dark matter or dark energy that seems to be required when general relativity is applied to describe the properties of the universe as a whole.

The remainder of this chapter argues that when searching for replacements for GR over short and long distances there is much to be learned from other branches of physics, where similar searches have revealed general constraints on how physics at different scales can relate to one another. The hard-won lessons learned there also have implications for gravitational physics, and this recognition is beginning to re-establish the connections between the gravitational and non-gravitational research communities.

In a nutshell, the lessons distilled from other areas of physics make it likely that it is much more difficult to modify gravity over very long distances than over very tiny ones. This is because very broad principles (like unitarity and stability) strongly restrict what is possible. The difficulty of modifying gravity over long distances is a very useful (but often neglected) clue when interpreting cosmological data, because it strongly constrains the theoretical options that are available. We ignore such clues at our peril.

This chapter is also meant to be colloquial rather than authoritative, and so citations are not thorough. My apologies to those whose work is not properly cited.

3.2 Modifying gravity over short distances

The demand to replace general relativity at short distances arises because quantum mechanics should make it impossible to have a spacetime description of geometry for arbitrarily small scales. For example, an accurate measurement of a geometry's curvature, R, requires positions to be measured with an accuracy, δ, smaller than the radius of curvature:

$$\delta^2 < 1/R\,. \tag{3.1}$$

For position measurements with resolution, δ, the uncertainty principle requires a momentum uncertainty, $p \simeq \hbar/\delta$, which implies an associated energy uncertainty, $E \simeq p\,c \simeq \hbar c/\delta$, or equivalently a mass $M \simeq E/c^2 \simeq \hbar/\delta c$. But the curvature associated with having this much energy within a distance of order δ is then $R \simeq GM/\delta^3 c^2 \simeq G\hbar/\delta^4 c^3 = \ell_p^2/\delta^4$, where ℓ_p defines the Planck length, $\ell_p^2 = G\hbar/c^3$, and G is Newton's constant. Requiring Eq. (3.1) then shows that there is a lower bound on the resolution with which spacetime can be measured:

$$\delta > \ell_p \simeq \sqrt{\frac{G\hbar}{c^3}} \simeq 1.6 \times 10^{-35}\,\mathrm{m}\,. \tag{3.2}$$

Although this is an extremely short distance (present experiments only reach down to about 10^{-19} m), it is also only a lower bound. Depending on how gravity really works over short distances, quantum gravity effects could arise at much longer scales.

Notice how crucial it is to this argument that the interaction strength, G, has dimensions of length squared (in fundamental units, for which $\hbar = c = 1$). Imagine performing a similar estimate for an electrostatic field. The coulomb interaction energy between two electrons separated by a distance δ is $E_c \simeq e^2/\delta$, where $q = -e$ denotes the electron's electric charge. But the energy required by the uncertainty principle to localize electrons this close to one another is $E \simeq \hbar c/\delta$, so the condition that this be smaller than E_c is

$$\alpha = \frac{e^2}{4\pi\hbar c} < 1\,, \tag{3.3}$$

where the fine-structure constant, $\alpha \simeq 1/137$, is dimensionless. This condition doesn't depend on δ because the relative strength of quantum fluctuations to electrostatic interactions does not change with distance.

3.2.1 Gravity and renormalizability

The observation that quantum fluctuations do not get worse at shorter distances in electrodynamics[1] but do for gravity can be more technically expressed as the statement that QED is a *renormalizable* quantum field theory (QFT) while GR is not. In QFT, small-distance quantum fluctuations appear (within perturbation theory) as divergences at small distances (or high momenta) when summing over all possible quantum intermediate states.

For instance, given a Hamiltonian, $H = H_0 + H_{\rm int}$, the second-order shift in the energy of a state $|n\rangle$ is

$$\delta_2 E_n = -\sum_m \frac{|\langle n|H_{\rm int}|m\rangle|^2}{E_m - E_n} \simeq -\int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{|\langle n|H_{\rm int}|\mathbf{p}\rangle|^2}{E(\mathbf{p}) - E_n} + \cdots, \tag{3.4}$$

where the approximate equality focuses on the sum over a basis of free single-particle states having energies $E(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}$ when performing the sum over $|m\rangle$. Because the combination $|\langle n|H_{\rm int}|\mathbf{p}\rangle|^2/[E(\mathbf{p}) - E_n]$ typically falls with large $p = |\mathbf{p}|$ like $1/p^3$ or slower, the integration over the momentum of the intermediate state diverges in the ultraviolet (UV), $p \to \infty$, limit. (Relativistic calculations organize these sums differently to preserve manifest Lorentz invariance at each step, but the upshot is the same.)

Renormalizability means that these divergences can all be absorbed into the unknown parameters of the theory – like the electron's charge and mass, for instance – whose values must in any case be inferred by comparison with experiments. As the above estimates suggest, the hallmark of a non-renormalizable theory is the appearance of couplings (like Newton's constant) having dimensions of length to a positive power (in fundamental units). Couplings like this ruin perturbative renormalizability because the more powers of them that appear in a result, the more divergent that result typically is.

For instance, a contribution that arises at nth order in Newton's constant usually depends on G through the dimensionless combination $(G\Lambda^2)^n \propto (\ell_p/\delta)^{2n}$, where $\Lambda \propto 1/\delta$ is the UV cutoff in momentum space (equivalently, δ is the small-distance

[1] There is a sense in which quantum effects in QED do get worse at smaller distances, because the theory is not asymptotically free. But this problem only arises logarithmically in δ, and so is much less severe than the power-law competition found above for gravity.

cutoff in position space). By contrast, having more powers of dimensionless couplings, or those having dimensions of inverse powers of length, do not worsen UV divergences. Ever-worsening divergences ruin the arguments which show for renormalizable theories that all calculations are finite once a basic set of couplings are appropriately redefined. Removal of divergences can be accomplished, but only by introducing an infinite number of coupling parameters to be renormalized.

Lack of renormalizability was for a long time regarded as a fundamental obstacle to performing any quantum calculations within gravity. After all, if every calculation is associated with a new parameter that absorbs the new divergences, whose value must be inferred experimentally, then there are as many parameters as observables and no predictions are possible. If this were really true, it would mean that any classical prediction of GR would come with incalculable theoretical errors due to the uncontrolled size of the quantum corrections. And the presence of such errors would render meaningless any detailed comparisons between classical predictions and observations, potentially ruining GR's observational successes. How can meaningful calculations be made?

3.2.2 Effective field theories

As it happens, tools for making meaningful quantum calculations using non-renormalizable theories exist, having been developed for situations where quantum effects are more important than they usually are for gravity [16, 30].

The key to understanding how to work with non-renormalizable theories is to recognize that they can arise as approximations to more fundamental, renormalizable physics, for which explicit calculations are possible. The way non-renormalizable theories arise in this case is as a low-energy/long-distance approximation in situations for which short-distance physics is unimportant, and so is coarse-grained or integrated out [17, 36].

For instance, consider the lagrangian density for the quantum electrodynamics of electrons and muons:

$$\mathcal{L}_{\rm QED} = -\frac{1}{4}\, F_{\mu\nu} F^{\mu\nu} - \overline{\psi}(\gamma^\mu D_\mu + m)\psi - \overline{\chi}(\gamma^\mu D_\mu + M)\chi\,, \tag{3.5}$$

where $m = m_e$ and ψ (or $M = m_\mu \gg m_e$ and χ) are the electron (or muon) mass and field. Here $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $D_\mu = \partial_\mu + ieA_\mu$, as usual, and γ^μ represents the Dirac matrices – that satisfy $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} = 2\text{diag}(-,+,+,+)$. This is a renormalizable theory because all parameters, e, m and M, have non-positive dimension when regarded as a power of length in fundamental units.

Suppose now we choose to examine observables only involving the electromagnetic interactions of electrons at energies $\omega \ll M$ (such as the energy levels of atoms, for instance). Muons should be largely irrelevant for these kinds of observables, but not completely so. Muons are not completely irrelevant because they can contribute to electron–photon processes at higher orders in perturbation theory as virtual states.

It happens that any such effects due to virtual muons can be described at low energies by the following *effective field theory* of electrons and photons only:

$$\begin{aligned}\mathcal{L}_{\rm eff} &= -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} - \overline{\psi}(\gamma^\mu D_\mu + m)\psi + \frac{k_1\,\alpha}{30\pi M^2} F^{\mu\nu}\Box F_{\mu\nu} + \cdots \\ &= -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} - \overline{\psi}(\gamma^\mu D_\mu + m)\psi + \frac{k_1\,\alpha}{15\pi M^2} (\overline{\psi}\gamma_\mu\psi)(\overline{\psi}\gamma^\mu\psi) + \cdots ,\end{aligned} \tag{3.6}$$

where the second line is obtained from the first by performing the field redefinition

$$A_\mu \to A_\mu + \frac{k_1\,\alpha}{15\pi M^2}\Big[\Box A_\mu - ie\,(\overline{\psi}\gamma_\mu\psi)\Big] + \cdots \tag{3.7}$$

In both equations the ellipses describe terms suppressed by more than two powers of $1/M$.

The lagrangian densities of Eqs. (3.5) and (3.6) are precisely equivalent in that they give precisely the same results for *all* low-energy electron/photon observables, provided one works only to leading order in $1/M^2$. If the accuracy of the agreement is to be at the one-loop level, then equivalence requires the choice $k_1 = 1$, and the effective interaction captures the leading effects of a muon loop in the vacuum polarization. If agreement is to be at the two-loop level, then $k_1 = 1 + \mathcal{O}(\alpha)$ captures effects coming from higher loops as well, and so on.

This example (and many many others) show that it must be possible to make sensible predictions using non-renormalizable theories. This must be so because the lagrangian of Eq. (3.6) is not renormalizable – its coupling has dimensions (length)^2 – yet it agrees precisely with the (very sensible) predictions of QED, Eq. (3.5). But it is important that this agreement only works up to order $1/M^2$.

If we work beyond order $1/M^2$ in this expansion, we can still find a lagrangian, $\mathcal{L}_{\rm eff}$, that captures all the effects of QED to the desired order. The corresponding lagrangian requires more terms than in Eq. (3.6), however, also including terms like

$$\mathcal{L}_4 = \frac{k_2\,\alpha^2}{90 M^4} (F_{\mu\nu}F^{\mu\nu})^2 , \tag{3.8}$$

that arise at order $1/M^4$. Agreement with QED in this case requires $k_2 = 1 + \mathcal{O}(\alpha)$. Sensible predictions can be extracted from non-renormalizable theories, but only if one is careful to work only to a fixed order in the $1/M$ expansion.

What is useful about this process is that an effective theory like Eq. (3.6) is much easier to use than is the full theory Eq. (3.5). And any observable whatsoever may be computed once the coefficients (k_1 and k_2 in the above examples) of the various non-renormalizable interactions are identified. This can be done by comparing its implications with those of the full theory for a few specific observables.

What about the UV divergences associated with these new effective interactions? They must be renormalized, and the many couplings required to perform this renormalization correspond to the many couplings that arise within the effective theory at successive orders in $1/M$. But predictiveness is not lost, because working to fixed order in $1/M$ means that only a fixed number of effective couplings are required in any given application.

At present this is the *only* known way to make sense of perturbatively non-renormalizable theories. In particular, it means that there is a hidden approximation involved in the use of a non-renormalizable theory – the low-energy, $1/M$, expansion – that may not have been obvious from the get-go.

3.2.3 GR as an effective theory

What would this picture mean if applied to GR? First, it would mean that GR must be regarded as the leading term in the low-energy/long-distance approximation to some more fundamental theory. Working beyond leading order would mean extending the Einstein–Hilbert action to include higher powers of curvatures and their derivatives, with the terms having the fewest derivatives being expected to dominate at low energies (for a review, see [6]).

Since we do not know what the underlying theory is, we cannot hope to compute the couplings in this effective theory from first principles as was done above for QED. Instead, we treat these couplings as phenomenological, ultimately to be determined from experiment.

The most general interactions involving the fewest curvatures and derivatives, that are consistent with general covariance, are:

$$\begin{aligned} -\frac{\mathcal{L}_{\rm eff}}{\sqrt{-g}} &= \lambda + \frac{M_p^2}{2}\, R + a_1\, R_{\mu\nu}\, R^{\mu\nu} \\ &\quad + a_2\, R^2 + a_3\, R_{\mu\nu\lambda\rho} R^{\mu\nu\lambda\rho} + a_4\, \Box R \\ &\quad + \frac{b_1}{m^2}\, R^3 + \frac{b_2}{m^2}\, R R_{\mu\nu} R^{\mu\nu} + \frac{b_3}{m^2}\, R_{\mu\nu} R^{\nu\lambda} R_\lambda{}^\mu + \cdots, \end{aligned} \tag{3.9}$$

where $R^\mu{}_{\nu\lambda\rho}$ is the metric's Riemann tensor, $R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}$ is its Ricci tensor, and $R = g^{\mu\nu} R_{\mu\nu}$ is the Ricci scalar, each of which involves precisely two derivatives of the metric.

The first term in Eq. (3.9) is the cosmological constant, which we drop because observations imply λ is (for some unknown reason, see below) extremely small.

Once this is done the leading term in the derivative expansion is the Einstein–Hilbert action whose coefficient, $M_p = (8\pi G)^{-1/2} = (\sqrt{8\pi}\,\ell_p)^{-1} \sim 10^{18}$ GeV, has dimensions of mass (when $\hbar = c = 1$), and is set by the value of Newton's constant. This is followed by curvature-squared terms having dimensionless effective couplings, a_i, and curvature-cubed terms with couplings inversely proportional to a mass, b_i/m^2 (not all of which are written in Eq. (3.9)).

Although the numerical value of M_p is known, the mass scale m appearing in the curvature-cubed (and higher) terms is not. But since it appears in the denominator it is the lowest mass scale to have been integrated out that should be expected to dominate. What its value should be depends on the scale of the applications one has in mind. For applications to the solar system or to astrophysics, m might reasonably be taken to be the electron mass, m_e. But for applications to inflation, where the scales of interest are much larger than m_e, m would instead be taken to be the lightest particle that is heavier than the scales of inflationary interest.

3.2.4 Power counting

The Einstein–Hilbert term should dominate at low energies (since it involves the fewest derivatives), and this expectation can be made more precise by systematically identifying which interactions contribute to a particular order in the semiclassical expansion. To do so we expand the metric about an asymptotically static background spacetime: $g_{\mu\nu} = \overline{g}_{\mu\nu} + 2h_{\mu\nu}/M_p$, and compute (say) the scattering amplitudes for asymptotic graviton states that impinge onto the geometry from afar.

If the energies, ω, of the incoming states are all comparable and similar to the curvatures scales of the background spacetime, dimensional analysis can be used to give an estimate for the energy dependence of an L-loop contribution to a scattering amplitude, $\mathcal{A}(\omega)$. Consider a contribution to this amplitude that involves E external lines and V_{id} vertices involving d derivatives and i attached graviton lines. Dimensional analysis leads to the estimate:

$$\mathcal{A}(\omega) \sim \omega^2 M_p^2 \left(\frac{1}{M_p}\right)^E \left(\frac{\omega}{4\pi M_p}\right)^{2L} \prod_i \prod_{d>2} \left[\frac{\omega^2}{M_p^2}\left(\frac{\omega}{m}\right)^{(d-4)}\right]^{V_{id}}. \tag{3.10}$$

Notice that no negative powers of ω appear here because general covariance requires derivatives come in pairs, so the index d in the product runs over $d = 4 + 2k$, with $k = 0, 1, 2, \ldots$

This last expression displays the low-energy approximation alluded to above because it shows that the small quantities controlling the perturbative expansion are ω/M_p and ω/m. Use of this expansion (and in particular its leading, classical limit – see below) presupposes both of these quantities to be small. Notice also that because $m \ll M_p$, factors of ω/m are much larger than factors of ω/M_p, but because they do not arise until curvature-cubed interactions are important, the perturbative expansion always starts off with powers of ω/M_p.

3.2.5 What justifies the classical approximation?

Equation (3.10) answers a question that is not asked often enough: What is the theoretical error made when treating gravitational physics in the classical approximation? What makes it so useful in this regard is that it quantifies the size of the contribution to $\mathcal{A}(\omega)$ (or other observables) arising both from quantum effects (i.e., loops with $L \geq 1$), and from terms normally not included in the lagrangian (such as higher-curvature terms). This allows an estimate of the size of the error that is made when such terms are not considered (as is often the case).

In particular, Eq. (3.10) justifies why classical calculations using GR work so well, and quantifies just how accurate their quantum corrections are expected to be. To see this, we ask which graphs dominate in the small-ω limit. For any fixed process (i.e., fixed E), Eq. (3.10) shows the dominant contributions are those for which

$$L = 0 \quad \text{and} \quad V_{id} = 0 \text{ for any } d > 2\,.$$

That is, the dominant contribution comes from arbitrary tree graphs constructed purely from the Einstein–Hilbert ($d = 2$) action. This is precisely the prediction of classical general relativity.

For instance, for the scattering of two gravitons about flat space, $g(p_1) + g(p_2) \to g(p_1') + g(p_2')$, we have $E = 4$, and Eq. (3.10) predicts the dominant energy-dependence to be $\mathcal{A}(\omega) \propto (\omega/M_p)^2$. This is borne out by explicit tree-level calculations [11], which give

$$\mathcal{A}_{\text{tree}} = 8\pi i G \left(\frac{s^3}{tu} \right), \tag{3.11}$$

for an appropriate choice of graviton polarizations. Here $s = -(p_1 + p_2)^2$, $t = (p_1 - p_1')^2$, and $u = (p_1 - p_2')^2$ are the usual Lorentz-invariant Mandelstam variables built from the initial and final particle 4-momenta, all of which are proportional to ω^2. This shows both that $\mathcal{A} \sim (\omega/M_p)^2$ to leading order, and that it is the physical, invariant, center-of-mass energy, ω_{cm}, that is the relevant scale against which m and M_p should be compared.

The next-to-leading contributions, according to Eq. (3.10), arise in one of two ways: either

$$L=1 \quad \text{and} \quad V_{id}=0 \text{ for any } d>2;$$

or

$$L=0, \sum_i V_{i4}=1, \quad \text{and} \quad V_{id}=0 \text{ for } d>4\,.$$

These correspond to one-loop (quantum) corrections computed only using Einstein gravity; plus a tree-level contribution including precisely one vertex from one of the curvature-squared interactions (in addition to any number of interactions from the Einstein–Hilbert term). The UV divergences arising in the first type of contribution are absorbed into the coefficients of the interactions appearing in the second type. Both are suppressed compared to the leading, classical, term by a factor of $(\omega/4\pi M_p)^2$. This estimate (plus logarithmic complications due to infrared divergences) is also borne out by explicit one-loop calculations about flat space [12, 13, 29].

This is the reasoning that shows why it makes sense to compute quantum effects, like Hawking radiation or inflationary fluctuations, within a gravitational context. For observables located a distance r away from a gravitating mass M, the leading quantum corrections are predicted to be of order $G\hbar/r^2c^3 = (\ell_p/r)^2$. For comparison, the size of classical relativistic corrections is set by $2GM/rc^2 = r_s/r$, where $r_s = 2GM/c^2$ denotes the Schwarzschild radius. At the surface of the Sun this makes relativistic corrections of order $GM_\odot/R_\odot c^2 \sim 10^{-6}$, while quantum corrections are $G\hbar/R_\odot^2 c^3 \sim 10^{-88}$. Clearly the classical approximation to GR is *extremely* good within the solar system.

On the other hand, although relativistic effects cannot be neglected near a black hole, since $2GM/r_s c^2 = 1$, the relative size of quantum corrections near the event horizon is

$$\left(\frac{\ell_p}{r_s}\right)^2 = \frac{G\hbar}{r_s^2 c^3} = \frac{\hbar c}{4GM^2}\,, \tag{3.12}$$

which is negligible provided $M \gg M_p$. Since M_p is of order tens of micrograms, this shows why quantum effects represent small perturbations for any astrophysical black holes,[2] but would not be under control for any attempt to interpret the gravitational field of an elementary particle (like an electron) as giving rise to a black hole.

[2] Small, but not irrelevant, since the decrease in mass predicted by Hawking radiation has no classical counterpart with which to compete.

3.2.6 Lessons learned

What do these considerations tell us about how gravity behaves over very small distances?

The good news is that it says that the observational successes of GR are remarkably robust against the details of whatever small-distance physics ultimately describes gravity over very small distances. This is because *any* microscopic physics that predicts the same symmetries (like Lorentz invariance) and particle content (a massless spin-2 particle, or equivalently a long-range force coupled to stress-energy) as GR, must be described by a generally covariant effective action like Eq. (3.9). Because this is dominated at low energies by the Einstein–Hilbert action, it suffices to get the low-energy particle content and symmetries right to get GR right in all of its glorious detail [10].

The bad news applies to those who think they know what the fundamental theory of quantum gravity really is at small scales, since whatever it is will be very hard to test experimentally. This is because all theories that get the bare minimum right (like a massless graviton), are likely to capture correctly all of the successes of GR in one fell swoop. At low energies the only difference between the predictions of *any* such theory is the value of the coefficients, a_i and b_i, etc. appearing in the low-energy lagrangian Eq. (3.9), none of which are yet observable.

There are two kinds of proposals that allow tests at low energies: those that change the low-energy degrees of freedom (such as by adding new light particles in addition to the graviton – more about these proposals below); and those that change the symmetries predicted for the low-energy theory. Prominent amongst this latter category are theories that postulate that gravity at short distances breaks Lorentz or rotational invariance, perhaps because spacetime becomes discrete at these scales.

At first sight, breaking Lorentz invariance at short distances seems batty, due to the high accuracy with which experimental tests verify the Lorentz-invariance of the vacuum within which we live. How could the world we see appear so Lorentz invariant if it is really not so deeper down? Surprisingly, experience with other areas of physics suggests this may not be so crazy an idea; we know of other, emergent, symmetries that can appear to be very accurate at long distances even though they are badly broken at short distances. Most notable among these is the symmetry responsible for conservation of baryon number, which has long been known to be an "accidental" symmetry of the Standard Model. This means that for *any* microscopic theory whose low-energy particle content is that of the SM, any violations of baryon number must necessarily be described by a non-renormalizable effective interaction [31, 34], and so be suppressed by a power of a large inverse mass, $1/M$. This suppression can be enough to agree with observations (like the absence of proton decay) if M is as large as 10^{16} GeV.

Could Lorentz invariance be similarly emergent? If so, it should be possible to find effective field theories for which Lorentz violation first arises suppressed by some power of a heavy scale, $1/M$, even if Lorentz invariance is not imposed from the outset as a symmetry of the theory. Unfortunately this seems hard to achieve, since in the absence of Lorentz invariance it is difficult[3] in an effective theory to explain why the effective terms

$$\partial_t \psi^* \partial_t \psi \qquad \text{and} \qquad \nabla\psi^* \cdot \nabla\psi \tag{3.13}$$

should have precisely the same coefficient in the low-energy theory. (See, however, [18] for some attempts.) The problem is that the coefficients of these terms are dimensionless in fundamental units, and so are unsuppressed by powers of $1/M$. But the relative normalization of these two terms governs the maximal speed of propagation of the corresponding particle, and there are extremely good bounds (for some particles better than a part in 10^{20}) on how much this can differ from the speed of light (see, for instance, [23] for a recent review).

This underlines why proponents of any particular quantum gravity proposal must work hard to provide the effective field theory (EFT) that describes their low-energy limit (see [21, 23] for some gravitational examples). Since all of the observational implications are contained within the effective theory, it is impossible to know without it whether or not the proposal satisfies all of the existing experimental tests. This is particularly true for proposals that claim to predict a few specific low-energy effects that are potentially observable (such as small violations of Lorentz invariance in cosmology). Even if the predicted effects should be observed, the theory must also be shown not to be in conflict with other relevant observations (such as the absence of Lorentz invariance elsewhere), and this usually requires an EFT formulation.

3.3 Modifying gravity over long distances

There has also been considerable activity over recent years investigating the possibility that GR might fail, but over very long distances rather than short ones. This possibility is driven most persuasively from cosmology, where the Hot Big Bang paradigm has survived a host of detailed observational tests, but only if the universe is pervaded by no less than *two* kinds of new exotic forms of matter: dark matter (at present making up $\sim 25\%$ of the universal energy density) and dark energy (comprising $\sim 70\%$ of the cosmic energy density). Because all of the evidence for the existence of these comes from their gravitational interactions, inferred

[3] The situation would be different in Euclidean signature, since then invariance under a lattice group of rotations can suffice to imply invariance under $O(4)$ transformations, at least for the kinetic terms.

using GR, the suspicion is that it might be more economical to interpret instead the cosmological tests as evidence that GR is failing over long distances.

But since the required modifications occur over long distances, their discussion is performed most efficiently within an effective lagrangian framework. These next paragraphs summarize my personal take on what has been learnt to this point.

3.3.1 Consistency issues

An important consideration when trying to modify gravity over long distances is the great difficulty in doing so in a consistent way. Almost all modifications so far proposed run into trouble with stability or unitarity, in that they predict unstable degrees of freedom like "ghosts," particles having negative kinetic energy. The presence of ghosts in a low-energy theory is generally regarded as poison because it implies there are instabilities. At the quantum level these instabilities usually undermine our understanding of particle physics and the very stability of the vacuum (see [9] for a calculation showing what can go wrong), but even at the classical level they typically ruin the agreement between the observed orbital decay of binary pulsars and GR predictions for their energy loss into gravitational waves.

The origin of these difficulties seems to be the strong consistency requirements that quantum mechanics and Lorentz invariance impose on theories of massless particles having spin 1 or higher [10, 28, 32], with static (non-derivative) interactions. A variety of studies indicate that a consistent description of particles with spins ≥ 1 always requires a local invariance, which in the case of spins 1, 3/2, and 2 corresponds to gauge invariance, supersymmetry, or general covariance, and this local symmetry strongly limits the kinds of interactions that are possible.[4] Although it remains an area of active research [14], at present the only systems known to satisfy these consistency constraints consist of relativistic theories of spins 0 through 1 coupled either to gravity or supergravity (possibly in more than four spacetime dimensions).

3.3.2 Dark matter

As might be expected, widespread acceptance of the existence of a hitherto-unknown form of matter requires the concordance of several independent lines of evidence, and this constrains one's options when formulating a theory for dark matter. It is useful to review this evidence when deciding whether it indicates a failure of GR or a new form of matter.

[4] The AdS/CFT correspondence [22] – a remarkable equivalence between asymptotically anti-de Sitter gravitational theories and non-gravitational systems in one lower dimensions – may provide a loophole to some of these arguments, although its ultimate impact is not yet known.

The evidence for dark matter comes from measuring the amount of matter in a region as indicated by how things gravitate towards it, and comparing the result with the amount of matter that is directly visible. Several types of independent comparisons consistently point to there being more than 10 times as much dark, gravitating material in space than is visible:[5]

- *Galaxies.* The total mass in a galaxy may be inferred from the orbital motion of stars and gas measured as a function of distance from the galactic center. The results, for large galaxies like the Milky Way, point to several times more matter than is directly visible.
- *Galaxy clusters.* Similar measurements using the motion of galaxies and temperature of hot gas in large galaxy clusters also indicate the presence of much more mass than is visible.
- *Structure formation.* Present-day galaxies and galaxy clusters formed through the gravitational amplification of initially small primordial density fluctuations. In this case the evidence for dark matter arises from the interplay of two facts: first, the initial density fluctuations are known to be very small, $\delta\rho/\rho \sim 10^{-5}$, at the time when the CMB (see below) was emitted; second, small initial fluctuations cannot be amplified by gravity until the epoch where non-relativistic matter begins to dominate the total energy density. But this does not give enough time for the initially small fluctuations to form galaxies unless there is much more matter present than can be accounted for by baryons. The amount required agrees with the amount inferred from the previous measures described above.

These in themselves do not show that the required dark matter need be exotic, the evidence for which also comes from several sources:

- *Primordial nucleosynthesis.* The total mass density of ordinary matter (baryons) in the universe can be inferred from the predicted relative abundance of primordial nuclei created within the Hot Big Bang. This predicted abundance agrees well with observations, and relies on the competition between nuclear reaction rates and the rate with which the universe cools. But both of these rates themselves depend on the net abundance of baryons in the universe: the nuclear reaction rates depend on the number of baryons present; and the cooling rate depends on how fast the universe expands, and so – at least in GR – on its total energy density. The success of the predictions of big bang nucleosynthesis (BBN) therefore fixes the fraction of the universal energy density which can consist of baryons, and implies that there can be at most a few times more baryons than what would be inferred by counting those that are directly visible.

[5] This is consistent with the cosmological evidence that dark matter is roughly 5 times more abundant than ordinary matter (baryons) because most of the ordinary matter is also dark, and so is also not visible.

- *The cosmic microwave background (CMB).* CMB photons provide an independent measure of the total baryon abundance. They do so because sound waves in the baryon density that are present when these photons were radiated are observable as small temperature fluctuations. Since the sound-wave properties depend on the density of baryons, a detailed understanding of the CMB temperature spectrum allows the total baryon density to be reconstructed. The result agrees with the BBN measure described above.

There are two main options for explaining these observations. Since dark matter is inferred gravitationally, perhaps the laws of gravity differ on extra-galactic scales from those in the solar system. Alternatively, there could exist a cosmic abundance of a new type of hitherto-undiscovered particle.

At present there are several reasons making it more likely that dark matter is explained by the presence of a new type of particle rather than by changing GR on long distances. First, as mentioned above, sensible modifications are difficult to make at long distances that lack ghosts and other inconsistencies. Second, no phenomenological modification of gravity has yet been proposed that accounts for all the independent lines of evidence given above (although there is a proposal that can explain the rotation of galaxies [24, 27]).

On the other hand, all that is required to obtain dark matter as a new form of matter is the existence of a new type of stable elementary particle having a mass and couplings similar to those of the Z boson, which is already known to exist. Z bosons would be excellent dark matter candidates if only they did not decay. A particle with mass and couplings like the Z boson, but which is stable – called a weakly interacting massive particle (WIMP) – would naturally have a relic thermal abundance in the Hot Big Bang that lies in the range observed for dark matter (for a review, see [15]). New particles with these properties are actually predicted by many current proposals for the new physics that is likely to replace the standard model at energies to be explored by the Large Hadron Collider (LHC).

At the present juncture the preponderance of evidence – the simplicity of the particle option and the difficulty of making a modification to GR that works – favors the interpretation of cosmological evidence as pointing to the existence of a new type of matter rather than a modification to the laws of gravity.

3.3.3 Dark energy

The evidence for dark energy is more recent, and incomplete, than that for dark matter. At present the evidence for its existence comes from two independent lines of argument:

- *Universal acceleration.* Since gravity is attractive, one expects an expanding universe containing only ordinary (and dark) matter and radiation to have a decelerating expansion rate. Evidence for dark energy comes from measurements indicating the universal expansion is *accelerating* rather than decelerating, obtained by measuring the brightness of distant supernovae [5, 25, 26]. According to GR, accelerated expansion implies the universe is dominated by something with an equation of state satisfying $p < -\rho/3$, which is not true for ordinary matter, radiation or dark matter.
- *Flatness of the universe.* An independent measure of the dark energy comes from the observed temperature fluctuations in the CMB. Because the CMB photons traverse the entire observable universe before reaching us, their properties on arrival depend on the geometry of the universe as a whole (and so also, according to GR, on its total energy density). Agreement with observations implies the total energy density is larger than the ordinary and dark matter abundances, which fall short by an amount consistent with the amount of dark energy required by the acceleration of the universe's expansion [20].

Again the theoretical options are the existence of a new form of energy density, or a modification of GR at long distances. Although there are phenomenological proposals for modifications that can cause the universe to accelerate (such as [14]), all of the previously described problems with long-distance modifications to GR also apply here.

By contrast, there is a very simple energy density that does the job, consisting simply of a cosmological constant – i.e. a constant $\lambda \simeq (3\times 10^{-3}\ \text{eV})^4$ in Eq. (3.9), for which $p = -\rho$. This is phenomenologically just what the doctor ordered, and agrees very well with the observations.

The theoretical difficulty here is that a cosmological constant is indistinguishable from the energy density of a Lorentz-invariant vacuum,[6] since both contribute to the stress tensor an amount $T_{\mu\nu} = \lambda\, g_{\mu\nu}$. In principle, this should be a good thing because we believe we can compute the vacuum energy. The problem is that ordinary particles (like the electron) contribute such an enormous amount – the electron gives $\delta\lambda \simeq m_e^4 \simeq (10^6\ \text{eV})^4$ – that agreement with the observed value requires a cancellation [33] to better than one part in 10^{36}.

3.3.4 Lessons learned

Dark matter and dark energy are two forms of exotic matter, whose existence is inferred purely from their gravitational influence on visible objects. It is tempting

[6] The only known loophole to this arises if extra dimensions exist, and are as large as 10 microns in size, because in this case the vacuum energy can be localized in the extra dimensions, and so curve these rather than the dimensions we see [1, 4, 8, 19]. Whether this, together with supersymmetry, can solve the problem is under active study [7].

to replace the need for two new things with a single modification to gravity over very large distances.

Yet the preponderance of evidence again argues against this point of view. First, it is difficult to modify GR at long distances without introducing pathologies. Second, it is difficult to find modifications that account for more than one of the several independent lines of evidence (particularly for dark matter). By contrast, it is not difficult to make models of dark matter (WIMPs) or dark energy (a cosmological constant). For dark energy this point of view runs up against the cosmological constant problem, which might indicate the presence of observably large extra dimensions, but for which no consensus yet exists.

3.4 Conclusions

In summary, modifications to general relativity are widely mooted over both large and small distances. This chapter argues that modifications at small distances are indeed very likely, and well worth seeking. But unless the modification takes place just beyond our present experimental reach ($\sim 10^{-19}$ m) [2, 3, 7], it is also likely to be very difficult to test experimentally. The basic obstruction is the decoupling from long distances of short-distance physics, a property most efficiently expressed using effective field theory methods. The good news is that this means that the many observational successes of GR are insensitive to the details of whatever the modification proves to be.

Modifications to GR over very long distances are also possible, and have been argued as more economical than requiring the existence of two types of unknown forms of matter (dark matter and dark energy). If so, consistency constraints seem to restrict the possibilities to supplementing GR by other very light spin-0 or spin-1 bosons (possibly in higher dimensions). The experimental implications of such modifications are themselves best explored using effective field theories. Unfortunately, no such modification has yet been found that accounts for all of the evidence for dark matter or energy in a way that is both consistent with other tests of GR and more economical than the proposals for dark matter or energy themselves.

To the extent that the utility of effective field theory relies on decoupling, one might ask: What evidence do we have that Planck-scale physics decouples? There are two lines of argument that bear on this question. First, once specific modifications to gravity are proposed it becomes possible to test whether decoupling takes place. Perhaps the best example of a consistent modification to gravity at short distances is string theory, and all the present evidence points to decoupling holding in this case. But more generally, if sub-Planckian scales do *not* decouple, one must ask: Why has science made progress at all? After all, although Nature comes to us with many scales, decoupling is what ensures we don't need to understand them all at once. If sub-Planckian physics does not decouple, what keeps it from appearing everywhere, and destroying our hard-won understanding of Nature?

Acknowledgments

I thank the editors for their kind invitation to contribute to this volume, and for their patience in awaiting my contribution. My understanding of this topic was learned from Steven Weinberg, who pioneered effective field theory techniques, and was among the first to connect the dots explicitly about gravity's interpretation as an effective field theory. My research is funded by the Natural Sciences and Engineering Research Council of Canada, as well as by funds from McMaster University and Perimeter Institute.

References

[1] Aghababaie, Y., Burgess, C. P., Parameswaran, S. L. and Quevedo, F. (2004) *Nucl. Phys. B* **680**, 389 [arXiv:hep-th/0304256].
[2] Antoniadis, I., Arkani-Hamed, N., Dimopoulos, S. and Dvali, G. (1998) *Phys. Lett. B* **436**, 257 [arXiv:hep-ph/9804398].
[3] Arkani-Hamed, N., Dimopoulos, S. and Dvali, G. (1998) *Phys. Lett. B* **429**, 263 [arXiv:hep-ph/9803315].
[4] Arkani-Hamed, N., Dimopoulos, S., Kaloper, N. and Sundrum, R. (2000) *Phys. Lett. B* **480**, 193, [hep-th/0001197].
[5] Bahcall, N., Ostriker, J.P., Perlmutter, S. and Steinhardt, P.J. (1999) *Science* **284**, 1481 [astro-ph/9906463].
[6] Burgess, C.P. (2004) *Living Rev. Rel.* **7**, 5 [gr-qc/0311082].
[7] Burgess, C.P. (2005) *AIP Conf. Proc.* **743**, 417 [arXiv:hep-th/0411140].
[8] Carroll, S.M. and Guica, M.M. (2003) [arXiv:hep-th/0302067].
[9] Cline, J.M., Jeon, S. and Moore, G.D. (2004) *Phys. Rev. D* **70**, 043543 [arXiv: hep-ph/0311312].
[10] Deser, S. (1970) *Gen. Rel. Grav.* **1**, 9 [arXiv:gr-qc/0411023].
[11] DeWitt, B.S. (1967) *Phys. Rev.* **162**, 1239.
[12] Donoghue, J.F. and Torma, T. (1999) *Phys. Rev.* **D60**, 024003 [hep-th/9901156].
[13] Dunbar, D.C. and Norridge, P.S. (1995) *Nucl. Phys.* **B433**, 181.
[14] Dvali, G., Gabadadze, G. and Porrati, M. (2000) *Phys. Lett.* **B485**, 208 [hep-th/0005016].
[15] Eidelman, S. *et al.* (2004) *Phys. Lett.* **B592**, 1.
[16] Gasser, G. and Leutwyler, H. (1984) *Ann. Phys. (NY)* **158**, 142.
[17] Gell-Mann, M. and Low, F.E. (1954) *Phys. Rev.* **95**, 1300.
[18] Groot Nibbelink, S. and Pospelov, M. (2005) *Phys. Rev. Lett.* **94**, 081601 [arXiv: hep-ph/0404271].
[19] Kachru, S., Schulz, M.B. and Silverstein, E. (2000) *Phys. Rev. D* **62**, 045021 [hep-th/0001206].
[20] Komatsu, E. *et al.* (2009), *ApJS*, **180**, 330 [arXiv:0803.0547].
[21] Kostelecky, V. A. (2004) *Phys. Rev. D* **69**, 105009 [arXiv:hep-th/0312310].
[22] Maldacena, J.M. (1998) *Adv. Theor. Math. Phys.* **2**, 231 [*Int. J. Theor. Phys.* **38** 1113] [arXiv:hep-th/9711200].
[23] Mattingly, D. (2005) *Living Rev. Rel.* **8**, 5 [arXiv:gr-qc/0502097].

[24] Milgrom, M. (1983) *Ap. J.* **270**, 365; 371; 384.
[25] Perlmutter, S. *et al.* (1997) *Ap. J.* **483**, 565 [astro-ph/9712212].
[26] Riess, A. G., *et al.* (1997) *Ast. J.* **116**, 1009 [astro-ph/9805201].
[27] Sanders, R. H. and McGaugh, S. S (2002) *Ann. Rev. Astron. Astrophys.* **40**, 263 [astro-ph/0204521].
[28] Weinberg, S. (1964) *Phys. Rev.* **134**, B882.
[29] Weinberg, S. (1965) *Phys. Rev.* **140**, 516.
[30] Weinberg, S. (1979a) *Physica* **96A**, 327.
[31] Weinberg, S. (1979b) *Phys. Rev. Lett.* **43**, 1566.
[32] Weinberg, S. and Witten, E. (1980) *Phys. Lett. B* **96**, 59.
[33] Weinberg, S. (1989) *Rev. Mod. Phys.* **61**, 1.
[34] Wilczek, F. and Zee, A. (1979) *Phys. Rev. Lett.* **43**, 1571.
[35] Will, C.M. (2001) *Living Rev. Rel.* **4**, 4 [gr-qc/0103036].
[36] Wilson, K.G. and Kogut, J.B. (1974) *Phys. Rept.* **12**, 75.

4

The small-scale structure of spacetime

STEVEN CARLIP

Several lines of evidence hint that quantum gravity at very small distances may be effectively two-dimensional. I summarize the evidence for such "spontaneous dimensional reduction," and suggest an additional argument coming from the strong-coupling limit of the Wheeler–DeWitt equation. If this description proves to be correct, it suggests a fascinating relationship between small-scale quantum spacetime and the behavior of cosmologies near an asymptotically silent singularity.

4.1 Introduction

Stephen Hawking and George Ellis prefaced their seminal book, *The Large Scale Structure of Space-Time*, with the explanation that their aim was to understand spacetime "on length-scales from 10^{-13} cm, the radius of an elementary particle, up to 10^{28} cm, the radius of the universe" [24]. While many deep questions remain, ranging from cosmic censorship to the actual topology of our universe, we now understand the basic structure of spacetime at these scales: to the best of our ability to measure such a thing, it behaves as a smooth (3+1)-dimensional Riemannian manifold.

At much smaller scales, on the other hand, the proper description is far less obvious. While clever experimentalists have managed to probe some features down to distances close to the Planck scale [43], for the most part we have neither direct observations nor a generally accepted theoretical framework for describing the very small-scale structure of spacetime. Indeed, it is not completely clear that "space" and "time" are even the appropriate categories for such a description.

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

But while a complete quantum theory of gravity remains elusive, we do have fragments: approximations, simple models, and pieces of what may eventually prove to be the correct theory. None of these fragments is reliable by itself, but when they agree with each other about some fundamental property of spacetime, we should consider the possibility that they are showing us something real. The thermodynamic properties of black holes, for example, appear so consistently that it is reasonable to suppose that they reflect an underlying statistical mechanics of quantum states.

Over the past several years, evidence for another basic feature of small-scale spacetime has been accumulating: it is becoming increasingly plausible that spacetime near the Planck scale is effectively two-dimensional. No single piece of evidence for this behavior is in itself very convincing, and most of the results are fairly new and tentative. But we now have hints from a number of independent calculations, based on different approaches to quantum gravity, that all point in the same direction. Here, I will summarize these clues, provide a further piece of evidence in the form of a strong-coupling approximation to the Wheeler–DeWitt equation, and discuss some possible implications.

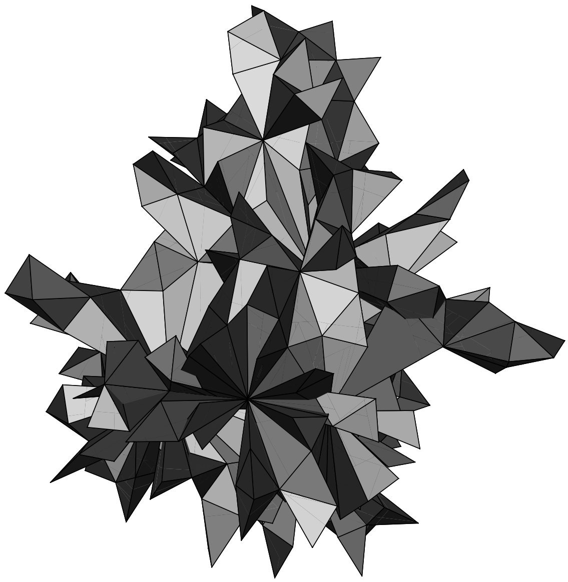
4.2 Spontaneous dimensional reduction?

Hints of short-distance "spontaneous dimensional reduction" in quantum gravity come from a number of places. Here I will review some of the highlights.

4.2.1 Causal dynamical triangulations

As we have learned from quantum chromodynamics – and from our colleagues in condensed matter physics – lattice approximations to the Feynman path integral can give us valuable information about the non-perturbative behavior of theories that may otherwise be extremely difficult to analyze. Lattice approximations to quantum gravity are not quite typical: in contrast to QCD, where fields live on a fixed lattice, gravity *is* the lattice [51], which forms a discrete approximation of a continuous spacetime geometry. Despite this difference, though, we might hope that a suitable lattice formulation could tell us something important about quantized spacetime.

The idea of combining Regge calculus with Monte Carlo methods to evaluate the gravitational path integral on a computer dates back to 1981 [54]. Until fairly recently, though, no good continuum limit could be found. Instead, the simulations typically yielded two unphysical phases, a "crumpled" phase with very high Hausdorff dimension and a two-dimensional "branched polymer" phase [41]. The causal dynamical triangulation program of Ambjørn, Jurkiewicz, and Loll [1–3] adds a



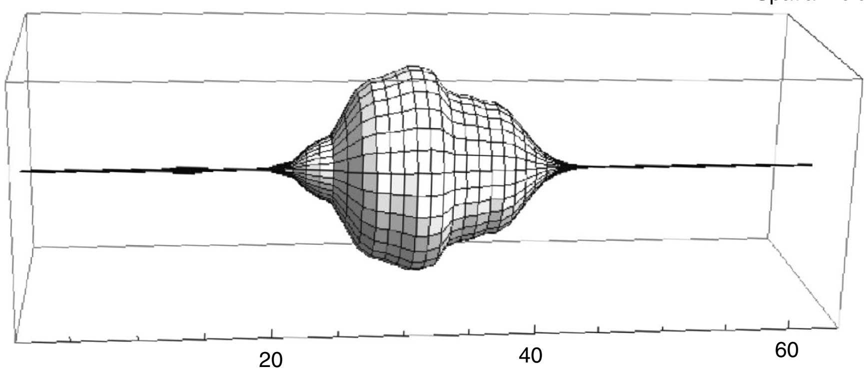


Figure 4.1 A spatial slice and a typical history contributing to the causal dynamical triangulations path integral
Source: R. Kommu and M. Sachs, UC Davis.

crucial new ingredient, a fixed causal structure in the form of a prescribed time-slicing. By controlling fluctuations in topology, this added structure suppresses the undesirable phases, and appears to lead to a good four-dimensional continuum picture. Results so far are very promising; in particular, the cosmological scale factor appears to have the correct semiclassical behavior [3, 4]. Figure 4.1 illustrates a typical time slice and a typical history contributing to the path integral in a simulation developed at UC Davis [37].

A crucial question for any such microscopic approximation is whether it can genuinely reproduce the four-dimensional structure we observe at “normal” distances. This is a subtle issue, which cannot be answered by merely looking at particular contributions to the path integral. As a first step, we need a definition of “dimension” for a discrete structure that may be very non-manifold-like at short distances. One natural choice – although by no means the only one – is the spectral dimension [5], the dimension as seen by a diffusion process or a random walker.

The basic idea of the spectral dimension is simple. In any structure on which a random walk can be defined, the associated diffusion process will gradually explore larger and larger regions of the structure. The more dimensions available for the random walk to explore, though, the longer this diffusion will take. Quantitatively, diffusion from an initial position x to a final position x' on a manifold M may be described by a heat kernel $K(x, x', s)$ satisfying

$$\left(\frac{\partial}{\partial s} - \Delta_x\right) K(x, x'; s) = 0, \qquad \text{with} \qquad K(x, x', 0) = \delta(x - x'), \tag{4.1}$$

where Δ_x is the Laplacian on M at x, and s is a measure of the diffusion time. Let $\sigma(x, x')$ be Synge's world function [58], one-half the square of the geodesic distance between x and x'. Then on a manifold of dimension d_S, the heat kernel generically behaves as

$$K(x, x'; s) \sim (4\pi s)^{-d_S/2} e^{-\sigma(x,x')/2s} (1 + \mathcal{O}(s)) \tag{4.2}$$

for small s. In particular, the return probability $K(x, x, s)$ is

$$K(x, x; s) \sim (4\pi s)^{-d_S/2}. \tag{4.3}$$

For any structure on which a diffusion process can be defined, be it a manifold or not, we can now use Eq. (4.3) to define an effective dimension d_S, the spectral dimension. On a lattice, in particular, we can determine the spectral dimension by directly simulating random walks. In the causal dynamical triangulation program, the ensemble average over the histories contributing to the path integral then gives a quantum spectral dimension. The results of such simulations yield a spectral dimension of $d_S = 4$ at large distances [3, 5]. This is a promising sign, indicating the recovery of four-dimensional behavior. Similarly, (2+1)-dimensional causal dynamical triangulations yield a large-distance spectral dimension of $d_S = 3$ [10, 37].

At short distances, though, the result is dramatically different. In both $3+1$ dimensions and $2+1$ dimensions, the small-scale spectral dimension falls to $d_S = 2$. This is the first, and perhaps the clearest, sign of spontaneous dimensional reduction at short distances.

As noted above, d_S is by no means the only definition of a generalized dimension, and one may worry about reading too much significance into this result. Note, though, that the heat kernel has a special significance in quantum field theory: propagators of quantum fields may be obtained as Laplace transforms of appropriate heat kernels. For a scalar field, in particular, the propagator is determined by the heat kernel (4.1), and the behavior of the spectral dimension implies a structure

$$G(x,x') \sim \int_0^\infty ds\, K(x,x';s) \sim \begin{cases} \sigma^{-2} & \text{at large distances} \\ \ln\sigma & \text{at small distances.} \end{cases} \tag{4.4}$$

The logarithmic short-distance behavior is the standard result for a two-dimensional conformal field theory. If one probes short distances with a quantum field, the field will thus act as if it lives in an effective dimension of two.

4.2.2 Renormalization group analysis

General relativity is non-renormalizable: conventional perturbative quantum field theory techniques yield an infinite number of higher-derivative counterterms, each with its own coupling constant. Nevertheless, the renormalization group *flow* of these coupling constants may provide us with valuable information about quantum gravity. In particular, a renormalization group analysis, even if incomplete or truncated, may offer a method for probing the theory at high energies and short distances.

One particularly dramatic possibility, first suggested by Weinberg [64], is that general relativity may be "asymptotically safe." Consider the full effective action for conventional gravity, with its infinitely many coupling constants. The renormalization group describes the dependence of these constants on energy scale, and in principle allows us to compute the high-energy/small-distance ("ultraviolet") couplings in terms of their low-energy/large-distance ("infrared") values. Under the renormalization group flow toward high energies, some of these constants may blow up, indicating that the description has broken down and that new physics is needed. An alternative possibility, though, is that the coupling constants might remain finite and flow to an ultraviolet fixed point. In that case, the theory would continue to make sense down to arbitrarily short distances. If, in addition, the critical surface – the space of such UV fixed points – were finite-dimensional, the long-distance coupling constants would be determined by a finite number of short-distance parameters: not quite renormalizability, but perhaps almost as good.

Distinguishing among these possibilities (and others) is extremely difficult, and we are still a long way from knowing whether quantum general relativity is asymptotically safe. But there is growing evidence for a UV fixed point, coming from various truncations of the effective action and from exact calculations in dimensionally reduced models [39, 47, 52]. For our present purposes, the key result of these calculations is that field operators acquire large anomalous dimensions; that is, under a change in mass scale, they scale differently than one would expect from dimensional analysis based on their classical "engineering dimension." In fact, these operators scale precisely as one would expect for the corresponding quantities in a *two-dimensional* field theory [52]. Moreover, the spectral dimension near the putative fixed point can be computed using field theoretical techniques, and the result is again $d_S = 2$ [38].

There is, in fact, a fairly general argument that if quantum gravity is asymptotically safe, it must be effectively two-dimensional at very short distances [47, 40]. Consider the dimensionless coupling constant $g_N(\mu) = G_N \mu^{d-2}$, where G_N is Newton's constant and μ is the mass scale that appears in the renormalization group flow. Under this flow,

$$\mu \frac{\partial g_N}{\partial \mu} = [d - 2 + \eta_N(g_N, \ldots)] g_N, \tag{4.5}$$

where the anomalous dimension η_N depends upon both g_N and any other dimensionless coupling constants in the theory. Evidently a free field (or "Gaussian") fixed point can occur at $g_N = 0$. For an additional non-Gaussian fixed point g_N^* to be present, though, the right-hand side of (4.5) must vanish: $\eta_N(g_N^*, \ldots) = 2 - d$.

But the momentum space propagator for a field with an anomalous dimension η_N has a momentum dependence $(p^2)^{-1+\eta_N/2}$. For $\eta_N = 2 - d$, this becomes p^{-d}, and the associated position space propagator depends logarithmically on distance. As I noted earlier, such a logarithmic dependence is characteristic of a two-dimensional conformal field. A variation of this argument shows that arbitrary matter fields interacting with gravity at a non-Gaussian fixed point exhibit a similar two-dimensional behavior [47].

4.2.3 Loop quantum gravity

A third hint of short-distance dimensional reduction comes from the area spectrum of loop quantum gravity [45]. States in this proposed quantum theory of gravity are given by spin networks, graphs with edges labeled by half-integers j (which represent holonomies of a connection), and vertices labeled by SU(2) intertwiners, that is, generalized Clebsch–Gordan coefficients. Given such a state, the area operator for a surface counts the number of spin network edges that puncture the surface, with each such edge contributing an amount

$$A_j \sim \ell_p^2 \sqrt{j(j+1)},$$

where

$$\ell_p = \sqrt{\frac{\hbar G}{c^3}} \tag{4.6}$$

is the Planck length.

While area and volume operators in loop quantum gravity are well understood, it has proven rather difficult to define a length operator. But since j is, crudely speaking, a quantum of area, one can (equally crudely) think of $\sqrt{j}$ as a sort of quantum of length. If we therefore define a length $\ell_j = \sqrt{j}\,\ell_p$, we can rewrite the spectrum as

$$A_j \sim \sqrt{\ell_j^2(\ell_j^2 + \ell_p^2)} \sim \begin{cases} \ell_j^2 & \text{for large areas} \\ \ell_p \ell_j & \text{for small areas.} \end{cases} \tag{4.7}$$

Like the propagator (4.4), this spectrum undergoes a change in scaling at small distances. Indeed, one can define a scale-dependent effective metric that reproduces the behavior (4.7), and use it to compute an effective spectral dimension. One again finds a dimension that decreases from four at large scales to two at small scales.

4.2.4 High-temperature strings

A fourth piece of evidence for small-scale dimensional reduction comes from the behavior of string theory at high temperatures. At a critical temperature, the Hagedorn temperature, the string theory partition function diverges, and the theory (probably) undergoes a phase transition. As early as 1988, Atick and Witten discovered that at temperatures far above the Hagedorn temperature, string theory has an unexpected thermodynamic behavior [6]: the free energy in a volume V varies with temperature as

$$F/VT \sim T. \tag{4.8}$$

For a field theory in d dimensions, in contrast, $F/VT \sim T^{d-1}$. So even though string theory lives in 10 or 26 dimensions, at high temperatures it behaves thermodynamically, as if spacetime were two-dimensional.

4.2.5 Anisotropic scaling models

A fifth sign of short-distance spontaneous dimensional reduction in quantum gravity comes from "Hořava–Lifshitz" models [29]. These are new models of gravity that exhibit anisotropic scaling, that is, invariance under constant rescalings

$$\mathbf{x} \to b\mathbf{x},\ t \to b^3 t.$$

This scaling property clearly violates Lorentz invariance, breaking the symmetry between space and time and picking out a preferred time coordinate. In fact, it is this breaking of Lorentz invariance that makes the models renormalizable: the field equations may now contain many spatial derivatives, leading to high inverse powers of spatial momentum in propagators that can tame loop integrals, while keeping only second time derivatives, thus avoiding negative energy states or negative norm ghosts. This might seem an unlikely way to quantize gravity – and indeed, these models may face serious low-energy problems [12] – but the hope is that Lorentz invariance might be recovered and conventional general relativity restored at low energies and large distances.

Hořava has calculated the spectral dimension in such models [30], and again finds $d_S = 2$ at high energies. In this case, the two-dimensional behavior can be traced to the fact that the propagators contain higher inverse powers of momentum; the logarithmic dependence on distance comes from integrals of the form

$$\int \frac{d^4p}{p^4} e^{ip\cdot(x-x')}.$$

Whether this is "real" dimensional reduction becomes a subtle matter, which may depend on how one operationally defines dimension. As before, though, the logarithmic behavior of propagators implies that quantum fields used to probe the structure of spacetime will act as if they are in a two-dimensional space.

4.2.6 Other hints

Hints of two-dimensional behavior come from several other places as well. In the causal set approach to quantum gravity, spacetime is taken to be fundamentally discrete, with a "geometry" determined by the causal relationships among points. In this setting, a natural definition of dimension is the Myrheim–Meyer dimension, which compares the number of causal relations between pairs of points to the corresponding number for points randomly sprinkled in a d_M-dimensional Minkowski space (see, for instance, [53]). For a small enough region of spacetime, one might guess that the causal structure is generic, coming from a random causal ordering. In that case, the Myrheim–Meyer dimension is approximately 2.38 – not quite 2, but surprisingly close [57].

Dimensional reduction has also appeared in an analysis of quantum field theory in a background "foam" of virtual black holes [17], although the effective dimension depends on the (unknown) black hole distribution function. A dimension of 2 also appears in Connes' non-commutative geometrical description of general relativity [16]; it is not clear to me whether this is related to the dimensional reduction considered here.

4.3 Strong coupling and small-scale structure

Let us suppose these hints are really telling us something fundamental about the small-scale structure of quantum gravity. We then face a rather bewildering question: which two dimensions? How can a four-dimensional theory with no background structure or preferred direction pick out two "special" dimensions at short distances? To try to answer this question, it is worth looking at one more approach to Planck-scale physics: the strong-coupling approximation to the Wheeler–DeWitt equation.

In the conventional Dirac quantization of canonical general relativity, the configuration space variables are given by the spatial metric g_{ij} on a time slice Σ, while their canonically conjugate momenta are related to the extrinsic curvature of the slice. The Hamiltonian constraint, which expresses invariance under diffeomorphisms that deform the slice Σ, then acts on states $\Psi[g]$ to give the Wheeler–DeWitt equation [18]

$$\left\{16\pi\ell_p^2 G_{ijkl}\frac{\delta}{\delta g_{ij}}\frac{\delta}{\delta g_{kl}} - \frac{1}{16\pi\ell_p^2}\sqrt{g}\,{}^{(3)}\!R\right\}\Psi[g]=0 \tag{4.9}$$

where

$$G_{ijkl}=\frac{1}{2}g^{-1/2}\left(g_{ik}g_{jl}+g_{il}g_{jk}-g_{ij}g_{kl}\right) \tag{4.10}$$

is the DeWitt metric on the space of metrics. The Wheeler–DeWitt equation is not terribly well-defined – the operator ordering is ambiguous, the product of functional derivatives must be regularized, and the wave functions $\Psi[g]$ should really be functions of spatial diffeomorphism classes of metrics – and it is not at all clear how to find an appropriate inner product on the space of solutions [14]. Nevertheless, the equation is widely accepted as a heuristic guide to the structure of quantum gravity.

Note now that spatial derivatives of the 3-metric appear only in the scalar curvature term ${}^{(3)}\!R$ in (4.9). As early as 1976, Isham [33] observed that this structure implied an interesting strong-coupling limit $\ell_p \to \infty$. As the Planck length becomes large – that is, as we probe scales near or below ℓ_p [48, 42] – the scalar curvature term becomes negligible. Neighboring spatial points thus effectively decouple, and the equation becomes ultralocal.

The absence of spatial derivatives greatly simplifies the Wheeler–DeWitt equation, which becomes exactly solvable. Its properties in this limit have been studied extensively [28, 32, 49, 50, 55, 60], and preliminary attempts have been made to restore the coupling between neighboring points by treating the scalar curvature term as a perturbation [21, 35, 42, 56]. The same kind of perturbative treatment of the scalar curvature is important classically in a very different setting: it is central to the Belinskii–Khalatnikov–Lifshitz (BKL) approach to cosmology near a spacelike singularity [8, 9, 25]. We can therefore look to this classical setting for clues about the small-scale structure of quantum gravity.

To understand the physics of the strong-coupling approximation, it is helpful to note that the Planck length (4.6) also depends on the speed of light. In fact, this approximation can also be viewed as a small c ("anti-Newtonian") approximation. As the Planck length becomes large, particle horizons shrink and light cones collapse to timelike lines, leading to the decoupling of neighboring points and the consequent ultralocal behavior [27]. In the completely decoupled limit, the classical solution at each point is a Kasner space,

$$ds^2 = dt^2 - t^{2p_1}dx^2 - t^{2p_2}dy^2 - t^{2p_3}dz^2$$
$$\left(-\frac{1}{3} < p_1 < 0 < p_2 < p_3, \quad p_1 + p_2 + p_3 = 1 = p_1^2 + p_2^2 + p_3^2\right). \tag{4.11}$$

More precisely – see, for example, [26] – the general solution is an arbitrary GL(3) transformation of a Kasner metric. This is still essentially a Kasner space, but now with arbitrary, not necessary orthogonal, axes.

For large but finite ℓ_p, the classical solution exhibits BKL behavior [8, 9, 25]. At any given point, the metric spends most of its time in a nearly Kasner form. But as the metric evolves, the scalar curvature can grow abruptly. The curvature term in the Hamiltonian constraint – the classical counterpart of the Wheeler–DeWitt equation – then acts as a potential wall, causing a Mixmaster-like "bounce" [44] to a new Kasner solution with different axes and exponents. In contrast to the ultralocal behavior at the strong-coupling limit, neighboring points are now no longer completely decoupled. But the Mixmaster bounces are chaotic [15], and the geometries at nearby points quickly become uncorrelated, with Kasner exponents occurring randomly with a known probability distribution [36].

We can now return to the problem of dimensional reduction. Consider a timelike geodesic in Kasner space, starting at $t = t_0$ with a random initial velocity. The geodesic equation is exactly integrable, and in the direction of decreasing t, the proper spatial distance traveled along each Kasner axis asymptotes to

$$\begin{aligned} s_x &\sim t^{p_1} \\ s_y &\sim 0 \\ s_z &\sim 0. \end{aligned} \tag{4.12}$$

Particle horizons thus shrink to lines, and geodesics effectively explore only one spatial dimension. In the direction of increasing t, the results are similar, though a bit less dramatic:

$$\begin{aligned} s_x &\sim t \\ s_y &\sim t^{\max(p_2, 1+p_1-p_2)} \\ s_z &\sim t^{p_3}. \end{aligned} \tag{4.13}$$

Since p_2, $1 + p_1 - p_2$, and p_3 are all less than one, a random geodesic again predominantly sees only one spatial dimension.

Since the heat kernel describes a random walk, one might expect this behavior of Kasner geodesics to be reflected in the spectral dimension. The exact form of heat kernel for Kasner space is not known, but Futamase [22] and Berkin [11] have evaluated $K(x, x'; s)$ in certain approximations. Both find behavior of the form

$$K(x, x; s) \sim \frac{1}{4\pi s^2}\left[1 + \frac{a}{t^2}s + \dots\right]. \tag{4.14}$$

The interpretation of this expression involves an order-of-limits question. For a fixed time t, one can always find s small enough that the first term in (4.14) dominates. This is not surprising: the heat kernel is a classical object, and we are still looking at a setting in which the underlying classical spacetime is four-dimensional. For a fixed return time s, on the other hand, one can always find a time t small enough that the second term dominates, leading to an effective spectral dimension of two. One might worry about the higher order terms in (4.14), which involve higher-inverse powers of t and might dominate at smaller times. But these terms do not contribute to the singular part of the propagator (4.4); rather, they give terms that go as positive powers of the geodesic distance, and are irrelevant for short-distance singularities, light cone behavior, and the like.

One can investigate the same problem via the Seeley–DeWitt expansion of the heat kernel [19, 23, 61],

$$K(x, x, s) \sim \frac{1}{4\pi s^2}\Big([a_0] + [a_1]s + [a_2]s^2 + \dots\Big). \tag{4.15}$$

The "Hamidew coefficient" $[a_1]$ is proportional to the scalar curvature, and vanishes for an exact vacuum solution of the field equations. In the presence of matter, however, the scalar curvature will typically increase as an inverse power of t as $t \to 0$ [8, 9]; this growth is slow enough to not disrupt the BKL behavior of the classical solutions near $t = 0$, but it will nevertheless give a diverging contribution to $[a_1]$.

This short-distance BKL behavior of the strongly coupled Wheeler–DeWitt equation may thus offer an explanation for the apparent dimensional reduction of quantum gravity at the Planck scale. We now have a possible answer to the question, "Which two dimensions?" If this picture is right, the dynamics picks out an essentially random timelike plane at each point. This choice, in turn, is reflected in the behavior of the heat kernel, and hence in the propagators and the consequent short-distance properties of quantum fields.

4.4 Spacetime foam?

The idea that the "effective infrared dimension" might differ from four goes back to work by Hu and O'Connor [31], but the relevance to short-distance quantum gravity was not fully appreciated at that time. To investigate this prospect further, though, we should better understand the physics underlying BKL behavior.

The BKL picture was originally developed as an attempt to understand the cosmology of the very early universe near an initial spacelike singularity. Near such a singularity, light rays are typically very strongly focused by the gravitational field, leading to the collapse of light cones and the shrinking of particle horizons. This "asymptotic silence" [25] is the key ingredient in the ultralocal behavior of the equations of motion, from which the rest of the BKL results follow.

Small-scale quantum gravity has no such spacelike singularity, so if a similar mechanism is at work, something else must account for the focusing of null geodesics. An obvious candidate is "spacetime foam," small-scale quantum fluctuations of geometry. Seeing whether such an explanation can work is very difficult; it will ultimately require that we understand the full quantum version of the Raychaudhuri equation. As a first step, though, let us start with the classical Raychaudhuri equation for the focusing of null geodesics,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \sigma_\alpha{}^\beta\sigma_\beta{}^\alpha + \omega_{\alpha\beta}\omega^{\alpha\beta} - R_{\alpha\beta}k^\alpha k^\beta. \tag{4.16}$$

Here, θ is the expansion of a bundle of light rays, essentially $(1/A)(dA/d\lambda)$ where A is the cross-sectional area; σ and ω are the shear and rotation of the bundle. Negative terms on the right-hand side of (4.16) decrease expansion, and thus focus null geodesics; positive terms contribute to defocusing.

If we naively treat (4.16) as an operator equation in the Heisenberg picture and take the expectation value, ignoring for the moment the need for renormalization, we see that quantum fluctuations in the expansion and shear should focus geodesics. Indeed,

$$\langle\theta^2\rangle = \langle\theta\rangle^2 + (\Delta\theta)^2, \tag{4.17}$$

with a similar equation for σ, so the uncertainties $\Delta\theta$ and $\Delta\sigma$ contribute negative terms to the right-hand side of (4.16). We can estimate the size of these fluctuations by noting that the expansion θ is roughly canonically conjugate to the cross-sectional area A [20]. Indeed, θ is the trace of an extrinsic curvature, which is, as usual, conjugate to the corresponding volume element. Keeping track of factors of $\hbar$ and G, we find an uncertainty relation of the form

$$\Delta\bar{\theta}\Delta A \sim \ell_p, \tag{4.18}$$

where $\bar{\theta}$ is the expansion averaged over a Planck distance along the congruence. In many approaches to quantum gravity – for instance, loop quantum gravity – we expect areas to be quantized in Planck units. It is thus plausible that $\Delta A \sim \ell_p^2$, which would imply fluctuations of θ of order $1/\ell_p$. This would mean very strong focusing at the Planck scale, as desired.

As I have presented it, this argument is certainly inadequate. To begin with, I have not specified which congruence of null geodesics I am considering. In the BKL analysis, a spacelike singularity determines a special null congruence. In short-distance quantum gravity, no such structure exists, and we will have to work hard to define θ as a genuine observable.

Moreover, while the classical shear contributes a negative term to the right-hand side of (4.16), the operator product $\sigma_\alpha{}^\beta\sigma_\beta{}^\alpha$ in a quantum theory must be renormalized, and need not remain positive. Indeed, it is known that this quantity becomes negative near the horizon of a black hole [13]; this is a necessary consequence of the fact that Hawking radiation decreases the horizon area. On the other hand, there are circumstances in which a particular average of this operator over a special null geodesic must be positive to avoid violations of the generalized second law of thermodynamics [63]. The question of whether quantum fluctuations and "spacetime foam" at the Planck scale can lead to something akin to asymptotic silence thus remains open.

4.5 What next?

In short, the proposal I am making is this: that spacetime foam strongly focuses geodesics at the Planck scale, leading to the BKL behavior predicted by the strongly coupled Wheeler–DeWitt equation. If this suggestion proves to be correct, it leads to a novel and interesting picture of the small-scale structure of spacetime. At each point, the dynamics picks out a "preferred" spatial direction, leading to approximately (1+1)-dimensional local physics. The preferred directions are presumably determined classically by initial conditions, but because of the chaotic behavior of BKL bounces, they are quickly randomized; in the quantum theory, they are picked out by an initial wave function, but again one expects evolution to scramble any initial choices. From point to point, these preferred directions vary continuously, but they oscillate rapidly [46]. Space at a fixed time is thus threaded by rapidly fluctuating lines, and spacetime by 2-surfaces; the leading behavior of the physics is described by an approximate dimensional reduction to these surfaces.

There is a danger here, of course: the process I have described breaks Lorentz invariance at the Planck scale, and even small violations at that scale can be magnified and lead to observable effects at large scales [43]. Note, though, that the symmetry violations in this scenario vary rapidly and essentially stochastically in both space and time. Such "non-systematic" Lorentz violations are harder to study, but there is evidence that they lead to much weaker observational constraints [7].

The scenario I have presented is still very speculative, but I believe it deserves further investigation. One avenue might be to use results from the eikonal approximation [34, 59, 62]. In this approximation, developed to study very high energy scattering, a similar dimensional reduction takes place, with drastically disparate time scales in two pairs of dimensions. Although the context is very different, the technology developed for this approximation could prove useful for the study of Planck scale gravity.

Acknowledgments

I would like to thank Beverly Berger, David Garfinkle, and Bei-Lok Hu for valuable advice on this project, and Leonardo Modesto and Lenny Susskind, among others, for asking hard questions. This work was supported in part by US Department of Energy grant DE-FG02-91ER40674.

References

[1] J. Ambjørn, J. Jurkiewicz, and R. Loll, *Phys. Rev. Lett.* **85**, 924 (2000), eprint hep-th/0002050.
[2] J. Ambjørn, J. Jurkiewicz, and R. Loll, *Phys. Rev. Lett.* **93**, 131301 (2004), eprint hep-th/0404156.
[3] J. Ambjørn, J. Jurkiewicz, and R. Loll, *Phys. Rev.* **D72**, 064014 (2005), eprint hep-th/0505154.
[4] J. Ambjørn, J. Jurkiewicz, and R. Loll, *Phys. Lett.* **B607**, 205 (2005), eprint hep-th/0411152.
[5] J. Ambjørn, J. Jurkiewicz, and R. Loll, *Phys. Rev. Lett.* **95**, 171301 (2005), eprint hep-th/0505113.
[6] J. J. Atick and E. Witten, *Nucl. Phys.* **B310**, 291 (1988).
[7] S. Basu and D. Mattingly, *Class. Quant. Grav.* **22**, 3029 (2005), eprint astro-ph/0501425.
[8] V. A. Belinskii, I. M. Khalatnikov, and E. M. Lifshitz, *Adv. Phys.* **19**, 525 (1970)
[9] V. A. Belinskii, I. M. Khalatnikov, and E. M. Lifshitz, *Adv. Phys.* **31**, 639 (1982).
[10] D. Benedetti and J. Henson, "Spectral geometry as a probe of quantum spacetime," eprint arXiv:0911.0401v1 [hep-th].
[11] A. L. Berkin, *Phys. Rev.* **D46**, 1551 (1992).
[12] D. Blas, O. Pujolas, and S. Sibiryakov, *JHEP* **0910**, 029 (2009), eprint arXiv:0906.3046 [hep-th].

[13] P. Candelas and D. W. Sciama, *Phys. Rev. Lett.* **38**, 1372 (1977).
[14] S. Carlip, *Rept. Prog. Phys.* **64**, 885 (2001), eprint gr-qc/0108040.
[15] D. F. Chernoff and J. D. Barrow, *Phys. Rev. Lett.* **50**, 134 (1983).
[16] A. Connes, "Noncommutative geometry year 2000," in *Highlights of Mathematical Physics*, edited by A. Fokas *et al.*, American Mathematical Society, 2002, eprint math.QA/0011193.
[17] L. Crane and L. Smolin, *Nucl. Phys.* **B267**, 714 (1986).
[18] B. S. DeWitt, *Phys. Rev.* **160**, 1113 (1967).
[19] B. S. DeWitt, "Dynamical theory of groups and fields," in *Relativity, Groups and Topology*, edited by C. DeWitt and B. DeWitt, Gordon and Breach, New York, 1964, pp. 587–820.
[20] R. Epp, "The symplectic structure of general relativity in the double null (2+2) formalism," eprint gr-qc/9511060.
[21] G. Francisco and M. Pilati, *Phys. Rev.* **D31**, 241 (1985).
[22] T. Futamase, *Phys. Rev.* **D29**, 2783 (1984).
[23] G. W. Gibbons, "Quantum field theory in curved spacetime," in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel, Cambridge University Press, Cambridge, 1979, pp. 639–79.
[24] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge, 1973.
[25] J. M. Heinzle, C. Uggla, and N. Röhr, *Adv. Theor. Math. Phys.* **13**, 293 (2009), eprint gr-qc/0702141.
[26] A. Helfer *et al.*, *Gen. Rel. Grav.* **20**, 875 (1988).
[27] M. Henneaux, *Bull. Math. Soc. Belg.* **31**, 47 (1979).
[28] M. Henneaux, M. Pilati, and C. Teitelboim, *Phys. Lett.* **110B**, 123 (1982).
[29] P. Hořava, *Phys. Rev.* **D79**, 084008 (2009), eprint arXiv:0901.3775 [hep-th].
[30] P. Hořava, *Phys. Rev. Lett.* **102**, 161301 (2009), eprint arXiv:0902.3657 [hep-th].
[31] B. L. Hu and D. J. O'Connor, *Phys. Rev.* **D34**, 2535 (1986).
[32] V. Husain, *Class. Quant. Grav.* **5**, 575 (1988).
[33] C. J. Isham, *Proc. R. Soc. London* **A351**, 209 (1976).
[34] D. Kabat and M. Ortiz, *Nucl. Phys.* **B388**, 570 (1992), eprint hep-th/9203082.
[35] A. A. Kirillov, *Int. J. Mod. Phys* **D3**, 431 (1994).
[36] A. A. Kirillov and G. Montani, *Phys. Rev.* **D56**, 6225 (1997).
[37] R. Kommu, in preparation.
[38] O. Lauscher and M. Reuter, *JHEP* **0510**, 050 (2005), eprint hep-th/0508202.
[39] D. F. Litim, *Phys. Rev. Lett.* **92**, 201301 (2004), eprint hep-th/0312114.
[40] D. F. Litim, *AIP Conf. Proc.* **841**, 322 (2006), eprint hep-th/0606044.
[41] R. Loll, *Living Rev. Relativity* **1**, 13 (1998), eprint gr-qc/9805049.
[42] K. Maeda and M. Sakamoto, *Phys. Rev.* **D54**, 1500 (1996), eprint hep-th/9604150.
[43] D. Mattingly, *Living Rev. Relativity* **8**, 5 (2005), eprint gr-qc/0502097.
[44] C. W. Misner, *Phys. Rev. Lett.* **22**, 1071 (1969).
[45] L. Modesto, "Fractal structure of loop quantum gravity," eprint arXiv:0812.2214 [gr-qc].
[46] G. Montani, *Class. Quant. Grav.* **12**, 2505 (1990).
[47] M. Niedermaier, *Class. Quant. Grav.* **24**, R171 (2007), eprint gr-qc/0610018.
[48] M. Pilati, "Strong coupling quantum gravity: an introduction," in *Quantum Structure of Space and Time*, edited by M. J. Duff and C. J. Isham, Cambridge University Press, Cambridge, 1982, pp. 53–69.
[49] M. Pilati, *Phys. Rev.* **D26**, 2645 (1982).
[50] M. Pilati, *Phys. Rev.* **D28**, 729 (1983).

[51] T. Regge, *Nuovo Cimento A* **19**, 558 (1961).
[52] M. Reuter and F. Saueressig, *Phys. Rev.* **D65**, 065016 (2002), eprint hep-th/0110054.
[53] D. Rideout, "Dynamics of causal sets," eprint arXiv:gr-qc/0212064.
[54] M. Rocek and R. M. Williams, *Phys. Lett.* **B104**, 31 (1981).
[55] C. Rovelli, *Phys. Rev.* **D35**, 2987 (1987).
[56] D. S. Salopek, *Class. Quant. Grav.* **15**, 1185 (1998), eprint gr-qc/9802025.
[57] R. Sorkin, personal communication.
[58] J. L. Synge, *Relativity: The General Theory*, North-Holland, Amsterdam, 1960.
[59] G. 't Hooft, *Phys. Lett.* **198B**, 61 (1987).
[60] C. Teitelboim, *Phys. Rev.* **D25**, 3159 (1982).
[61] D. V. Vassilevich, *Phys. Rept.* **388**, 279 (2003), eprint hep-th/0306138.
[62] H. L. Verlinde and E. P. Verlinde, *Nucl. Phys.* **B371**, 246 (1992), eprint hep-th/9110017.
[63] A. Wall, "Proving the achronal averaged null energy condition from the generalized second law," eprint arXiv:0910.5751.
[64] S. Weinberg, "Ultraviolet divergences in quantum theories of gravitation," in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel, Cambridge University Press, Cambridge, 1979, pp. 790–831.

5

Ultraviolet divergences in supersymmetric theories

KELLOG STELLE

Recent advances in calculational techniques permit the study of ultraviolet structure in maximal super Yang–Mills and maximal supergravity theories at heretofore unattainable loop orders. Hints from string theory suggest that maximal supergravity might have a similar ultraviolet behavior in $D = 4$ spacetime dimensions as maximal super Yang–Mills theory and so be ultraviolet convergent. However, what is known of field theoretic non-renormalization theorems suggests only that $\frac{1}{2}$-BPS counterterms are excluded. A key test of the relative finiteness properties of the two theories will be the ultraviolet divergences in $D = 5$ maximal supergravity at the four-loop level. This chapter constitutes a review of the arguments that lead to these remarkable results.

5.1 Introduction

Obtaining an acceptable quantum theory of gravity is the key remaining problem in fundamental theoretical physics. A basic difficulty in formulating such a theory was already recognized in the earliest approaches to this problem in the 1930s: the dimensional character of Newton's constant gives rise to ultraviolet divergent quantum correction integrals. Naïve power counting of the degree of divergence Δ of an L-loop diagram in D-dimensional gravity theories yields the result

$$\Delta = (D-2)L + 2, \tag{5.1}$$

which grows linearly with loop order, implying the requirement for higher and higher-dimensional counterterms to renormalize the divergences. In the 1970s, this was confirmed explicitly in the first Feynman diagram calculations of the radiative

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

corrections to systems containing gravity plus matter [51]. The time lag between the general perception of the UV divergence problem and its first concrete demonstration was due to the complexity of Feynman diagram calculations involving gravity. The necessary techniques were an outgrowth of the long struggle to control, consistently with Lorentz invariance, the quantization of non-abelian Yang–Mills theories in the standard model of weak and electromagnetic interactions and likewise in quantum chromodynamics.

With the advent of supergravity [23, 29] in the mid-1970s, hopes rose that the specific combinations of quantum fields present in supergravity theories might possibly tame the gravitational UV divergence problem. Indeed, it turns out that all the irreducible supergravity theories in four-dimensional spacetime, i.e. theories in which all fields are irreducibly linked to gravity by supersymmetry transformations, have remarkable cancellations in Feynman diagrams at the one- and two-loop levels.

There is a sequence of such irreducible (or "pure") supergravity models, characterized by the number N of local, that is, spacetime-dependent, spinor transformation parameters. In four-dimensional spacetime, minimal, i.e. $N = 1$, supergravity thus has four supersymmetries corresponding to the components of a single Majorana spinor transformation parameter. The maximal possible supergravity [21] in four-dimensional spacetime has $N = 8$ spinor parameters, i.e. 32 independent supersymmetries.

The hopes for "miraculous" UV divergence cancellations in supergravity were subsequently dampened, however, by the realization that the divergence-killing powers of supersymmetry most likely do not extend beyond the two-loop order for generic pure supergravity theories [24, 36, 38, 47]. The three-loop invariant that was anticipated is quartic in curvatures, and has a purely gravitational part given by the square of the Bel–Robinson tensor [24].

In the 1980s and 1990s, the problem of quantum gravity took on a different character with the flowering of superstring theory. In this, the UV divergence problems of gravity are resolved by a completely different mechanism. The basic field-theory point-particle states are replaced by relativistic extended-object states and the changes in propagator structure cause quantum amplitudes to become ultraviolet convergent. These string theory developments pushed the UV divergence properties of point-particle field theories out of the limelight, leaving the UV problems of supergravity itself in an unclear state.

Early studies [24, 36, 38, 47] indicated that non-renormalization theorems deriving from superspace quantization might allow the first UV divergences at the loop orders shown in Table 5.1 for various spacetime dimensions. These divergence expectations involved combining a number of different requirements. These included the fraction of a given theory's supersymmetry that could be linearly (or "off-shell") realized in the Feynman rules, in respect of gauge invariances and the

Table 5.1. *Early expectations for maximal supergravity first divergences, assuming half the supercharges (i.e., 16) are linearly realizable*

Dimension D	11	10	8	7	6	5	4
Loop order L	2	2	1	2	3	2	3
General form	$\partial^{12}R^4$	$\partial^{10}R^4$	R^4	$\partial^4 R^4$	$\partial^6 R^4$	R^4	R^4

Table 5.2. *Early expectations for maximal super Yang–Mills first divergences, assuming half the supercharges (i.e., 8) are linearly realizable*

Dimension D	10	8	7	6	5	4
Loop order L	1	1	2	3	4	∞
General form	$\partial^2 F^4$	F^4	$\partial^2 F^4$	$\partial^2 F^4$	F^4	finite

availability of an invariant with respect to the full supersymmetry which remains non-vanishing subject to the classical equations of motion. (Only such divergences require counterterms; divergence structures that vanish subject to the classical field equations can be removed by field redefinitions.)

Despite the bleak UV outlook, a faint hope persisted among some researchers that the maximal $N = 8$ supergravity might have very special UV properties, in distinction to the non-maximal cases. This hope was bolstered by the complete ultraviolet finiteness in $D = 4$ dimensions of maximal $N = 4$ supersymmetric Yang–Mills theory [19, 37, 49]. This was the first interacting UV–finite theory in four spacetime dimensions. In higher dimensional spacetimes, super Yang–Mills theory itself becomes divergent, however, with the expected first divergence orders providing a very useful comparison to the supergravity cases. Based upon similar analysis to the supergravity cases, early expectations were that the initial divergences in super Yang–Mills theory would occur for various spacetime dimensions as shown in Table 5.2.

Of course, one cannot prove the presence of an ultraviolet divergence simply by studying non-renormalization theorems. This requires a proper calculation. There is a "folk expectation," however, that in complicated Feynman diagram calculations, vanishing results do not happen without a clear underlying reason. But this does not strictly rule out the possibility of "miraculous" cancellations which are not predicted by the non-renormalization theorems. But the prevailing expectation runs against the occurrence of such "miracles."

It is just such apparently "miraculous" UV divergence cancellations that have recently been confirmed, however, in remarkable 3-loop and 4-loop calculations in maximal super Yang–Mills [6] and maximal supergravity [7, 10]. Performing such calculations at high loop orders requires a departure from textbook Feynman-diagram methods [4], because the standard approaches can produce astronomical numbers of terms. Instead of following the standard propagator and vertex methods for supergravity calculations, Bern *et al.* used another technique which dates back to Feynman: loop calculations can be performed using the unitarity properties of the quantum S-matrix. These involve cutting rules which reduce higher-loop diagrams to sums of products of leading-order "tree" diagrams without internal loops. This use of unitarity is an outgrowth of the optical theorem in quantum mechanics for the imaginary part of the S-matrix.

In order to obtain information about the real part of the S-matrix, an additional necessary element in the unitarity-based technique is the extended use of dimensional regularization to render UV-divergent diagrams formally finite. In dimensional regularization, the dimensionality of spacetime is changed from 4 to $4-\varepsilon$, where ε is a small adjustable parameter. Traditional Feynman diagram calculations also often use dimensional regularization, but normally one just focuses on the leading $1/\varepsilon$ poles in order to subtract them in a renormalization program. In the unitarity-based approach, all orders in ε need to be retained. This gives rise to logarithms in which real and imaginary contributions are related.

In maximal $N=8$ supergravity theory, the complexity of the quantum amplitudes factorizes, with structures involving the particular field types occurring on the external legs of an amplitude multiplying a much simpler set of scalar-field Feynman diagrams. It is to the latter that the unitarity-based methods may be applied. Earlier applications [5] of the cutting-rule unitarity methods based on iteration of two-particle cuts gave an expectation that one might have cancellations for $D < 10/L+2$, where D is the spacetime dimension and L is the number of Feynman diagram loops (for $L > 1$). Already, this gave an expectation that $D=4$ maximal supergravity might have cancellations of the UV divergences at the $L=3$ and $L=4$ loop orders. This would leave the next significant test at $L=5$ loops. In the ordinary Feynman-diagram approach, a full calculation at this level would involve something like 10^{30} terms. Even using the unitarity-based methods, such a calculation would be a daunting, but perhaps not impossible, task.

The impressive new elements in the 3- and 4-loop calculations of Bern *et al.* are the completeness of the calculations and the unexpected further patterns of cancellations found. This might suggest a possibility of UV cancellations at yet

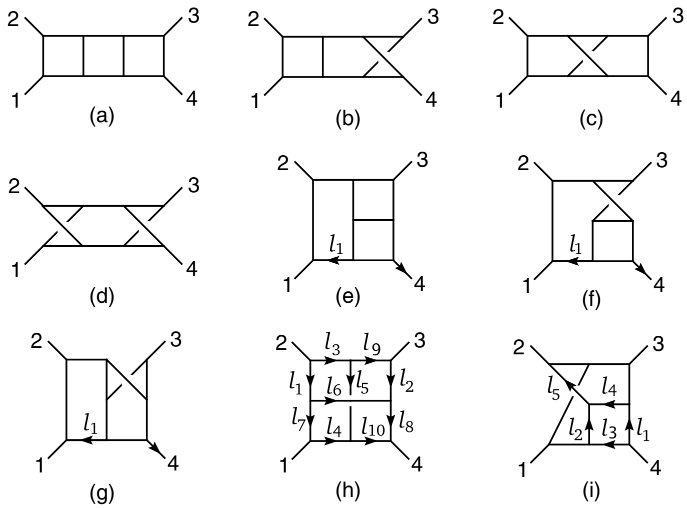


Figure 5.1 3-Loop Feynman diagram types leading to ultraviolet finiteness of maximal supergravity at this loop order. Diagrams (a)–(g) can be analyzed using iterated 2-particle cuts, leading to an expectation of ultraviolet divergence cancellation. Diagrams (h) and (i) cannot be treated this way, but the result of summing all diagrams (a)–(i) is an even deeper cancellation of the leading UV behavior than anticipated.

higher loop orders. Although the various 3-loop diagram classes were already individually expected to be finite on the basis of earlier work by Bern *et al.*, the new results show that the remaining finite amplitudes display additional cancellations, rendering them "superfinite." In particular, the earlier work employed iterated 2-particle cuts and did not consider all diagram types. The new complete calculations display further cancellations between diagrams that can be analyzed using iterated 2-particle cuts and additional diagrams that cannot be treated in this way. The nine 3-loop diagram types are shown in Figure 5.1. The end result is that the sum of all diagram types is more convergent by two powers of external momentum than might otherwise have been anticipated. This work was subsequently [8] reformulated in a way that makes the UV properties manifest diagram-by-diagram.

At the 4-loop level [10], there are 50 diagram topologies and the cancellations are more remarkable still. Keeping track of the combinatorics required computerization of most of the calculation. In the end, the 50 structures combine to yield finite results in both $D=4$ and $D=5$.

Does such a mechanism cascade in higher-order diagrams, rendering the maximal $N=8$ theory completely free of ultraviolet divergences? No one knows at present. Such a scenario might pose puzzling questions for the superstring program, where it has been assumed that ordinary supergravity theories need string ultraviolet completions in order to form consistent quantum theories. On the one hand, there are hints [12, 33] from superstring theory that precisely such an all-orders divergence cancellation might take place in the $N=8$ theory. On the other hand, it is not clear exactly what one can learn from superstring theory about purely perturbative field-theory divergences. Examples in Kaluza–Klein theories underline the likely non-commutativity between quantization and truncation of infinite sets of massive states. So, even though supergravity may be obtained as a zero-slope limit of superstring theory at the tree level, it is not clear what one can learn about quantized supergravity from superstring theory.

One thing that seems clear is that ordinary Feynman-diagram techniques coupled with the non-renormalization theorems of supersymmetry are unlikely to be able to explain finiteness properties of $N=8$ supergravity at arbitrary loop order. The earlier expectations [24, 36, 38, 47] were that the first divergences unremovable by field redefinitions would occur at three loops in all pure $D=4$ supergravities. A key element in this earlier anticipation was the expectation that the maximal amount of supersymmetry that can be linearly realized in Feynman diagram calculations (aka "off-shell supersymmetry") is half the full supersymmetry of the theory, or 16 out of 32 supercharges for the maximal $N=8$ theory. The precise choice of the set of half the supersymmetry generators which are to be linearly realized does not make an important difference to the result. Thus, for instance, light-cone methods [49, 19, 48] employ a non-Lorentz-covariant quantization method which, however, maintains manifestly the automorphism symmetry (SU(4) for $N=4$ super Yang–Mills, SU(8) for $N=8$ supergravity). Although the intermediate steps are quite different, the allowed counterterms are in the end the same – although translation between formalisms can be somewhat involved.

Similarly to the way in which chiral integrals of $N=1$, $D=4$ supersymmetry achieve invariance in integrals over less than the theory's full superspace, provided the integrand satisfies a corresponding BPS-type constraint, there are analogous invariants involving integration over varying portions of an extended supersymmetric theory's full superspace [36]. "$\frac{1}{2}$-BPS" operators require integration over just half the full set of fermionic θ coordinates. And if half the full supersymmetry were the maximal amount that is linearly realizable (with corresponding constraints on the diagrams from the corresponding Ward identities), such operators would be the first to be allowed as UV counterterms [40]. The detailed analysis of which counterterms are allowed and which are not involves the extent to which counterterm integrands must be manifestly gauge invariant and must also respect the other rigid automorphism symmetries of a theory. For this purpose, the background field method [36] is a very useful tool. It can be used to calculate effective action contributions with only background fields on external lines, and yet one can use Ward identities for the background–quantum split [39] to show how to renormalize all diagrams occurring at higher orders. In the case of supersymmetric gauge theories, this implies that counterterm integrands must be written in terms of background gauge connections, and must not involve prepotentials explicitly. This analysis of counterterm eligibility was backed up by agreement with the results of explicit Feynman diagram computations in maximal super Yang–Mills theory [50].

A main point of contention in the early analyses concerned the eligibility of the $\frac{1}{2}$-BPS supergravity counterterms, either at $L=2$, $D=5$ or at $L=3$, $D=4$. These have power-counting weight $\Delta=8$, and have generic structure $(\text{curvature})^4$. The detailed structure at leading quartic order in fields of this supergravity counterterm in maximal supergravity [36] reveals its similarity to the analogous $(\text{field strength})^4$ candidate super Yang–Mills counterterm. Written in terms of $D=4$ on-shell linearized superfields, the candidate counterterms are written in terms of the basic scalar superfields carrying totally antisymmetric R-symmetry indices ϕ_{ij} carrying a **6** of SU(4) in maximal super Yang–Mills or W_{ijkl} carrying a **70** of SU(8) in maximal supergravity:

$$
\begin{aligned}
\Delta I_{\text{SYM}} &= \int d^4x (d^4\theta d^4\bar{\theta})_{\mathbf{105}}\, Tr(\phi^4)_{\mathbf{105}} && \mathbf{105} \leftrightarrow \\
\Delta I_{\text{SG}} &= \int d^4x (d^8\theta d^8\bar{\theta})_{\mathbf{232848}} (W^4)_{\mathbf{232848}} && \mathbf{232848} \leftrightarrow
\end{aligned}
\tag{5.2}
$$

These structures reveal explicitly their $\frac{1}{2}$-BPS character, involving integrations over just 8 of the 16 odd superspace coordinates for maximal super Yang–Mills and 16 of the 32 odd coordinates for maximal supergravity. Despite these structures seemingly in accord with Feynman rules linearly realizing half the full supersymmetry, there were early indications from the unitarity-based calculations (using iterated 2-particle cuts) that, at least in the maximal super Yang–Mills case at $D=5$, $L=4$ [5], the divergences might nonetheless cancel.

The results of [7] showed definitely that the expectation of $\frac{1}{2}$-BPS operators as the first allowed maximal supergravity and super Yang–Mills counterterms is not sufficiently restrictive. Counterterm analysis quickly came up with an update, however. In [3, 17] it was shown that there exist non-Lorentz-covariant off-shell formulations with half supersymmetry *plus one*, i.e., 9 supercharges for maximal super Yang–Mills and 17 supercharges for maximal supergravity. The super Yang–Mills formulation dimensionally reduces down to (8,1) supersymmetry in $D=2$. In $D=2$ maximal supergravity, there is an analogous formulation with (16,1) supersymmetry. Although a full analysis of this formalism in spacetime dimensions $D>2$ has not been completed, the existence of such a formalism would be just enough to push out the boundary of the non-renormalization theorems so that the

Table 5.3. *Maximal supergravity divergence expectations based on "$\frac{1}{2}$ supersymmetry + 1" non-renormalization theorems*

Dimension D	11	10	8	7	6	5	4
Loop order L	2	2	1	2	3	4	5
BPS degree	0	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
General form	$\partial^{12}R^4$	$\partial^{10}R^4$	R^4	$\partial^4 R^4$	$\partial^6 R^4$	$\partial^6 R^4$	$\partial^4 R^4$

$\frac{1}{2}$-BPS counterterms (5.2) are just *ruled out* instead of just being allowed. With this enhanced understanding of the off-shell possibilities, the expectations in various dimensions for the first maximal supergravity divergences would be as in Table 5.3.

These anticipated enhanced non-renormalization theorems would dispel some of the mystery of the "miraculous" cancellations found in [7, 10], but this cannot be the full story. The 3-loop cancellations in $D=4$ maximal supergravity would thereby be explained as normal, albeit recondite, consequences of supersymmetry. The 4-loop cancellations in $D=4$ maximal supergravity found in [10] would similarly be explained. But a different sort of problem poses itself at the 4-loop level in $D=5$. As one can see from (5.1), $D=6$, $L=3$ and $D=5$, $L=4$ yield the same naïve degree of divergence: $\Delta=14$. In dimensional regularization, one only sees divergent expressions corresponding to logarithmic divergences in a standard momentum cutoff regularization. The naïve degree of divergence thus corresponds to the power counting weight of the fields on the external lines plus the number of derivatives/momenta on those lines. A $\Delta=14$ counterterm for a 4-point operator (the first on-shell non-vanishing structure) corresponds generically to $\partial^6 R^4$ in *both* $D=6$, $L=3$ and $D=5$, $L=4$. The $D=6$, $L=3$ divergence is known to occur [7], while the $D=5$, $L=4$ structure is now known *not* to occur [10]. The required $D=5$ counterterm might be expected to be formed simply by the dimensional reduction of the corresponding $D=6$ structure. How can such apparently similar structures differ in eligibility under a non-renormalization theorem?

One possibility might be a dimensional dependence on the possible degree of off-shell linearly realized supersymmetry. Maximal supersymmetric Yang–Mills theory in dimensions $D\geq 4$ might perhaps be a model for what is happening. In $D=4$, there exists a more extensive off-shell formulation of maximal super Yang–Mills using "harmonic superspace" methods [22, 30]. These realize, at the cost of an infinite number of auxiliary fields, 12 off-shell supersymmetries out of the total 16 supersymmetries of maximal super Yang–Mills theory. It is not known whether there are maximal supergravity analogues of this harmonic superspace formalism, but if there existed one with 28 off-shell supersymmetries in $D=6$ while such a formalism somehow failed to exist for $D=5$, this could explain the calculational results. But this seems like a tall order: why should it become more difficult to realize supersymmetries off-shell in lower dimensions?

A possible clarification of this issue follows again from maximal super Yang–Mills theory: there exists an analogue of the above $D=6 \leftrightarrow D=5$ supergravity divergence conundrum. In the super Yang–Mills case, one needs to go beyond the leading "single-trace" gauge-group structure to see this, however. As one can see from Table 5.2, divergences of power counting weight $\Delta=10$, i.e. of generic structure $\partial^2 F^4$, are expected both at ($L=2$, $D=7$) and ($L=3$, $D=6$). Indeed, the ($L=2$, $D=7$) were found by explicit Feynman diagram calculation long ago [50]. These $D=7$ results revealed both single-trace $\mathrm{tr}(\partial^2 F^4)$ and double-trace $(\partial \mathrm{tr} F^2)^2$ divergences. Exactly similar structures were expected in the ($L=3$, $D=6$) case. These 3-loop results were only recently obtained, however, using unitarity calculational methods in [9]. The result was surprising: although the single-trace operators occur as expected, the double-trace operators are mysteriously absent for ($L=3$, $D=6$).

5.2 Algebraic renormalization and ectoplasm

There is another approach to the non-renormalization theorems. This employs "algebraic renormalization" using BRST cohomological techniques and it has now been used to give yet another demonstration of the finiteness of $D=4$, $N=4$ super Yang–Mills [1]. In this approach, one first uses the Callan–Symanzik equation to relate the renormalizations of the classical action for a given theory to gamma functions corresponding to renormalization of the Lagrangian taken as an operator insertion into the quantum amplitudes. Then, supersymmetry covariance gives rise to a cohomological analysis of the counterterm integrands (written in the first instance not in superspace but in ordinary spacetime, with constituent fields taken on-shell). Consistency requires that the cocycle structure for the classical Lagrangian agrees with that for an allowed counterterm. For $\frac{1}{2}$-BPS counterterms, it turns out that the cocycle structures do not agree with that of the classical Lagrangian, either in maximal super Yang–Mills (thus once again ruling out F^4 counterterms) or in maximal supergravity (thus once again ruling out the R^4 counterterm) [17].

The algebraic approach to the renormalization of maximally supersymmetric theories was discussed in some detail in [17, 18]. Here we present a brief synopsis of the method. The basic idea is to study the symmetry properties of the effective action Γ algebraically. In the absence of a convenient superfield formalism or a set of auxiliary fields needed to linearly realize the supersymmetry, one can decide to discard them completely and to work with the full set of symmetries acting on the ordinary component fields. The supersymmetry transformations are then non-linear, however, and the algebra only closes modulo gauge transformations and the equations of motion, and gauge-fixing is not manifestly supersymmetric. Nonetheless, all of these technical problems can be overcome by use of the Batalin–Vilkoviski (BV) version of standard BRST techniques [2, 25, 41]. An important point is that one needs to introduce supersymmetry ghosts, which are constant anticommuting spinors in the case of rigid supersymmetry. After some algebra, one can then show that there is a BRST operator s under which any putative counterterm should be invariant. If we express such an invariant as an integral of a spacetime D-form, $\mathcal{L}_D$ say, then one has

$$s\mathcal{L}_D + d_0\mathcal{L}_{D-1,1} = 0 \tag{5.3}$$

where $\mathcal{L}_{D-1,1}$ is a spacetime $(D-1)$-form with ghost number one and d_0 is the spacetime exterior derivative. Applying s to (5.3) and using the facts that it is nilpotent and anticommutes with d_0, one deduces that

$$s\mathcal{L}_{D-1,1} + d_0\mathcal{L}_{D-2,2} = 0 \tag{5.4}$$

and so on. Thus we obtain a cocycle of the operator $\delta := s + d_0$ whose components $\mathcal{L}_{D-q,q}$ are $(D-q)$-forms with ghost number q. Now the question of whether a given invariant is allowed as a counterterm can be reformulated in terms of the anomalous dimension of the same invariant when considered as a composite operator insertion. Furthermore, one can include sources for all of the terms in the invariant's cocycle, to be considered as operator insertions; this can also be done for the original starting Lagrangian. One then concludes that an invariant will be allowable as a counterterm if it has the same cocycle structure as that of the initial Lagrangian. This is the algebraic supersymmetry non-renormalization theorem [17].

The invariants of interest are, in particular, gauge-invariant, and on such operators s essentially reduces to the supersymmetry BRST operator.[1] So a term $\mathcal{L}_{D-q,q}$ in a cocycle corresponds to a $(D-q)$-form with q additional spinor indices which have to be totally symmetrized since the supersymmetry ghost is a commuting object. This implies that the cocycle is equivalent to a closed D-form in superspace. We can therefore study the possible solutions to the algebraic non-renormalization problem systematically using superspace cohomology. This is advantageous because it allows us to study the problem starting at the lowest dimension and working upwards rather than the other way round. Since the top component has many terms

[1] There is also a translational ghost whose rôle is related to that of the t_0 operator as discussed below.

besides the leading bosonic one, this can be a rather complicated object to construct. Of course, any invariant can also be presented as a superspace integral, and in general a superfield integrand will have many more components than those appearing in the cocycle, so it seems that the algebraic approach implies that the essential part of a superfield integral is actually the part that appears in the closed super D-form. For example, as we shall see, the cocycle associated with a $\frac{1}{2}$-BPS invariant is actually longer than the cocycle for the classical Lagrangian, whereas the cocycle associated with a full superspace integral is the same as that for the Lagrangian. In this way, one sees that $\frac{1}{2}$-BPS operators are in fact ruled out as counterterms, while full superspace integrals are allowed, as expected.

It has been known for some time, moreover, that one can write a supersymmetric invariant as a spacetime integral in terms of a closed superform, a procedure which has been dubbed "ectoplasm" [32]. Suppose M is a supermanifold with a D-dimensional body (i.e., purely bosonic subspace) M_0 and suppose L_D is a closed D-form on M. The formula for a supersymmetric invariant I is

$$I = \int_{M_0} L_{D,0}(x, \theta = 0), \tag{5.5}$$

where $L_{D,0}$ is the purely bosonic component of L_D with respect to some coordinate basis and where (x, θ) are (even, odd) coordinates on M. It is easy to see that this does give a supersymmetry invariant because, under an infinitesimal diffeomorphism of M, a closed form changes by a total derivative, and a spacetime supersymmetry transformation is given by the leading term of an odd superdiffeomorphism in its θ-expansion. Since an exact D-form integrates to zero, it follows that we need to analyze the Dth de Rham cohomology group of M in order to find the possible invariants. This has nothing to do with topology, however, since the forms we are interested in have components which are gauge-invariant functions of the physical fields and this leads to non-trivial cohomology even for flat supermanifolds.

We next give a brief review of some of the essential aspects of superspace cohomology. We shall only consider flat superspace here, as appropriate for the super Yang–Mills case with which we shall be mostly concerned. The standard superinvariant basis 1-forms are

$$\begin{aligned} E^a &= dx^a - \frac{1}{2} d\theta^\alpha (\Gamma^a)_{\alpha\beta} \theta^\beta \\ E^\alpha &= d\theta^\alpha, \end{aligned} \tag{5.6}$$

which are dual to the usual invariant derivatives (∂_a, D_α). As we are going to focus on maximal super Yang–Mills, the index α can be thought of as a 16-component

$D=10$ chiral spinor index, although in $D<10$ it will stand for a combined spinor and R-symmetry index. Similarly, Γ^a denotes the 10-dimensional gamma matrices which reduce to a direct product of internal and spinor matrices.

The fact that the tangent spaces of a superspace (even in the curved case) split invariantly into even and odd subspaces implies that one can introduce a bi-grading on the spaces Ω^n of differential n-forms, $\Omega^n = \oplus_{p+q=n}\Omega^{p,q}$. We can also split the exterior derivative d into the following components with the indicated bidegrees [15]

$$d = d_0(1,0) + d_1(0,1) + t_0(-1,2). \tag{5.7}$$

In a general superspace there is also a component t_1 of bidegree $(2,-1)$ but it vanishes in flat space (and does not play a crucial cohomological role in any case). $d_0 = E^a\partial_a$ and $d_1 = E^\alpha D_\alpha$ are respectively even and odd exterior derivatives, while t_0 is an algebraic operation involving the dimension zero torsion, which is proportional to Γ. For $\omega \in \Omega^{p,q}$,

$$(t_0\omega)_{a_2\ldots a_p\beta_1\ldots\beta_{q+2}} \sim (\Gamma^{a_1})_{(\beta_1\beta_2}\omega_{a_1\ldots a_p\beta_3\ldots\beta_{q+2})}. \tag{5.8}$$

Since $d^2=0$ we find, amongst other relations,

$$t_0^2 = 0 \tag{5.9}$$

$$t_0 d_1 + d_1 t_0 = 0 \tag{5.10}$$

$$d_1^2 + t_0 d_0 + d_0 t_0 = 0. \tag{5.11}$$

Equation (5.9) implies that we can define t_0-cohomology groups $H_t^{p,q}$ [15]. We can then define a new odd derivative d_s acting on elements of these groups by

$$d_s[\omega] := [d_1\omega], \tag{5.12}$$

where $\omega \in [\omega] \in H_t^{p,q}$, with $[\omega]$ denoting the cohomology class of a t_0-closed form ω. Equations (5.10) and (5.11) then imply that these definitions are independent of the choice of representative ω and that $d_s^2=0$. This means that we can define the so-called spinorial cohomology groups $H_s^{p,q}$ [20, 46]. The point of these definitions is that they enable us to solve for the superspace cohomology of d in terms of the spinorial cohomology groups. Specifically, suppose that the lowest-dimensional non-zero component (i.e., the one with the largest number of odd indices) of some closed D-form L_D is $L_{D-q,q}$, for some q. Then, since $dL_D=0$, we have $t_0 L_{D-q,q}=0$, and since we are interested in cohomology, the starting component will correspond to an element of $H_t^{D-q,q}$. The next component of $dL_D=0$ then tells us that $d_s[L_{D-q,q}]=0$. Thereafter, if one can solve this equation, one can solve for all of the higher components of L_D in terms of $L_{D-q,q}$.[2] There may, of course, be other solutions to the problem with lowest components of different bidegrees, but this is precisely what is needed for there to be non-trivial examples of non-renormalization theorems, as this implies the existence of more than one type of cocycle. Another important consideration is that any putative lowest component of a closed D-form must lead to a non-zero $L_{D,0}$.

Next, consider for example the cohomology of $N=1, D=10$ superspace (see [13] for more details). Interestingly enough, this turns out to be closely related to the pure spinor approach to supersymmetry [42, 43]. Consider first $H_t^{0,q}$. Let $\omega \in \Omega^{0,q}$ and let $\bar{\omega} := u^{\alpha_q} \dots u^{\alpha_1} \omega_{\alpha_1 \dots \alpha_q}$ where u is a (commuting) pure spinor, $u\Gamma^a u = 0$. Clearly, if $\omega \mapsto \omega + t_0 l$, where $l \in \Omega^{1,q-2}$, $\bar{\omega}$ is unchanged, so that $H_t^{0,q}$ is isomorphic to the space of q-fold pure spinors appearing in pure spinor cohomology [11]. The t_0 cohomology groups for $1 \le p \le 5$ are again spaces of pure spinor-type objects but with additional antisymmetrized vector indices. This arises because of the gamma-matrix identities which are responsible for the kappa-symmetry of the string and fivebrane actions. In form notation these are

$$t_0 \Gamma_{1,2} = t_0 \Gamma_{5,2} = 0, \tag{5.13}$$

where $\Gamma_{p,2}$ denotes a symmetric gamma-matrix with p even indices viewed as a $(p,2)$-form. For our problem, only the second of these relations is relevant. For example, suppose $\omega \in \Omega^{3,q}$ can be written

$$\omega_{3,q} = \Gamma_{5,2} \lambda^2_{,q-2}, \tag{5.14}$$

where the notation indicates that two of the even indices on $\Gamma_{5,2}$ are to be contracted with the two vector indices on λ. Then it is clearly the case that ω is t_0-closed. Furthermore, in cohomology, the object λ can be taken to be of pure spinor type on its odd indices. Constructions such as this are not possible for $p \ge 6$, and it turns out that all such t_0-cohomology groups vanish.

Although it would seem that there are quite a lot of cohomology groups available which one might consider as possible lowest components for Lagrangian D-forms, it turns out that there is only one type of cocycle in $N=1, D=10$, with lowest component $L_{5,5}$ [13]. This is due to the fact that this is the only case which can lead to a non-trivial $L_{10,0}$. So any closed Lagrangian form in $D=10$ has a lowest component of the form

$$L_{5,5} = \Gamma_{5,2} M_{0,3}, \tag{5.15}$$

[2] In principle there can be higher-order obstructions but these do not arise in the examples discussed here.

where $d_s[M_{0,3}] = 0$. The simplest example of this is for an unconstrained scalar superfield S, which corresponds to a full superspace integral

$$M_{\alpha\beta\gamma} = T_{\alpha\beta\gamma,\delta_1\ldots\delta_5} D^{11\delta_1\ldots\delta_5} S, \tag{5.16}$$

where T is an invariant tensor constructed from gamma-matrices [11] and $D^{11\alpha_1\ldots\alpha_5}$ is the dual of the antisymmetrized product of 11 D_αs. The tensor T is symmetric on $\alpha\beta\gamma$ and totally antisymmetric on the δs. Now, a closed D-form in D dimensions gives rise to a closed $(D-1)$-form in $(D-1)$ dimensions under dimensional reduction, so this means that we can immediately construct the corresponding cocycle associated with any non-BPS invariant in dimensions $4 < D < 10$; it will have lowest component $L_{D-5,5} \sim \Gamma_{D-5,2} M_{0,3}$, where $\Gamma_{D-5,2}$ is the dimensional reduction of $\Gamma_{5,2}$.

The next example we shall consider is the (on-shell) Lagrangian. This is an example of a Chern–Simons (CS) invariant. In D dimensions, such an invariant can be constructed starting from a closed, gauge-invariant $(D+1)$-form $W_{D+1} = dZ_D$, where Z_D is a potential D-form [44], provided that it has the property of Weil triviality [16], i.e., provided it can also be written as dK_D for some gauge-invariant D-form K_D. If this is true, then $L_D := K_D - Z_D$ is closed and can be used to construct an integral invariant via the ectoplasm formula. For the $D = 10$ SYM action, the appropriate W_{11} is $H_7 Tr(F^2)$, where, in flat superspace, the closed 7-form $H_7 \sim \Gamma_{5,2}$. This 11-form W_{11} is easily seen to have the correct property, with the lowest component of K_D being $K_{8,2}$; Z can be chosen to be $H_7 Q_3$, where Q_3 is the SYM Chern–Simons 3-form, $dQ_3 = Tr(F^2)$. The lowest non-zero component of L_{10} is then

$$L_{5,5} = -\Gamma_{5,2} Q_{0,3}. \tag{5.17}$$

We can again reduce this formula to any dimension $4 < D < 10$, and so conclude that the lowest term in the Lagrangian form for all of these cases is $L_{D-5,5} = -\Gamma_{D-5,2} Q_{0,3}$. We are therefore able to conclude that the cocycle type of the Lagrangian is the same as that of a non-BPS invariant in all dimensions $D \geq 5$, and therefore that full superspace integrals are not protected by the algebraic non-renormalization theorem. This is, of course, as expected since the dawn of the subject.

We now move on to discuss BPS invariants, starting with the $\frac{1}{2}$-BPS case. There are two of these corresponding to single- and double-trace F^4 invariants, corresponding to the originally anticipated SYM divergences for ($L = 4$, $D = 5$). There is not a lot of difference between the two group-invariant structures from the point of view of superspace cohomology, so we shall focus on the double-trace as it will be useful in the subsequent discussion of the double-trace $\frac{1}{4}$-BPS case. In $D = 10$, this invariant is again of CS type with $W_{11} = H_3 F^4$, but the CS nature is lost for $D \le 8$ owing to the low rank of $H_3 \sim \Gamma_{1,2}$ and so we shall derive the Lagrangian form starting from scratch in $D = 7$ and below.[3]

In $D \le 8$, the SYM field strength multiplet is a scalar superfield W_r, $r = 1,\ldots,n = 10 - D$, whose independent components are the physical scalars and spinors together with the spacetime field strength. From it one can construct two bilinear multiplets of relevance: the Konishi multiplet $K := Tr(W_r W_r)$ and the supercurrent, $J_{rs} := Tr(W_r W_s) - \frac{1}{n}\delta_{rs} K$. The supercurrent is itself a $\frac{1}{2}$-BPS operator, but it is ultra-short in the sense that its θ-expansion only goes up to θ^4 (as opposed to θ^8 for a standard $\frac{1}{2}$-BPS superfield). It has 128+128 components, while the Konishi superfield is an unconstrained scalar superfield in the interacting theory. The supercurrent contains all of the conserved currents of super Yang–Mills theory: the R-currents, the supersymmetry current, the energy–momentum tensor and an identically conserved topological current for the gauge fields.

From quadratic expressions in J, one obtains scalar superfields in various representations of the R-symmetry group. The totally symmetric, traceless representation is the $\frac{1}{2}$-BPS multiplet we are interested in. Let us consider first the $D = 7$ case. We can take the R-symmetry group to be SU(2) and use i, j etc. to denote doublet SU(2) indices. The supercurrent is J_{ijkl}, while the one-half BPS multiplet is $B_{i_1\ldots i_8} := J_{(i_1\ldots i_4} J_{i_5\ldots i_8)}$. It obeys the constraint

$$D_{\alpha i} B_{j_1\ldots j_8} = \varepsilon_{i(j_1} \mathcal{L}_{\alpha j_2\ldots j_8)}, \tag{5.18}$$

where the spinor index can take on eight values. The lowest component of the Lagrangian 7-form is an $L_{0,7}$ of the form

$$L_{\alpha_1 i_1,\ldots,\alpha_7 i_7} = \eta_{(\alpha_1\alpha_2} \ldots \eta_{\alpha_5\alpha_6} \mathcal{L}_{\alpha_7) i_1\ldots i_7}, \tag{5.19}$$

where $\eta_{\alpha\beta}$ is the (symmetric) charge-conjugation matrix. It is straightforward to verify that this defines an element of $H_s^{0,7}$ and that it contains a singlet $L_{7,0}$, which yields the spacetime double-trace F^4 invariant. Furthermore, it is not difficult to show that this 7-form cannot be brought to the same form as that of the Lagrangian by the addition of an exact term. This shows that the $\frac{1}{2}$-BPS invariant has a different cocycle structure to that of the Lagrangian, and thus would be ruled out as an allowed counterterm structure. Admittedly, this fact is not directly relevant in $D = 7$ as this counterterm cannot arise there anyway for dimensional reasons, as the naïve power counting would lead one to expect a first divergence of power-counting weight $\Delta = 10$, as one can see in Table 5.2.

[3] F^4 arises at one loop in $D = 8$ where it is divergent; this is compatible with both algebraic and superspace non-renormalization theorems because they are not valid at one loop.

Now let us consider the $\frac{1}{4}$-BPS $\Delta = 10$ double-trace invariant d^2F^4. It turns out that this can be written as a subsuperspace integral of an associated pseudo-$\frac{1}{2}$-BPS superfield which is constructed from the one above by the insertion of two contracted spacetime derivatives, one on each factor of J. This allows one to write down a candidate Lagrangian form immediately with lowest component given as in (5.19) but where now $\mathcal{L} \sim \partial\chi \cdot \partial\mathcal{J}$, where $D\mathcal{J} \sim \chi$. In this case, however, one can show that

$$L_{0,7} = d_1 K_{0,6} + t_0 K_{1,5} \tag{5.20}$$

for some $K_{0,6}$ and $K_{1,5}$ which are constructed explicitly in terms of bilinears in the components of J. This is enough to show that this Lagrangian form is cohomologically equivalent to the Lagrangian form since there are only two types of cocycle in $D = 7$. Hence the double-trace $\frac{1}{4}$-BPS invariant in $D = 7$ is not protected.

The above Lagrangian form can be reduced straightforwardly to give a closed 6-form in $D = 6$ which must also have the same cocycle structure as the action. One might, therefore, be tempted to conclude that this invariant cannot be protected in $D = 6$ either. However, the R-symmetry group in $D = 7$ is SU(2) while for $N = 2$, $D = 6$ it is $SU(2) \times SU(2)$ and there is no guarantee that the reduced 6-form will have the full R-symmetry. For this reason we need to analyze $N = 2$, $D = 6$ starting again from the supercurrent.

The $N = 2$, $D = 6$ supersymmetry algebra is

$$\begin{aligned}
\{D_{\alpha i}, D_{\beta j}\} &= i\varepsilon_{ij}(\gamma^a)_{\alpha\beta}\partial_a \\
\{D^{\alpha i'}, D^{\beta j'}\} &= i\varepsilon^{i'j'}(\gamma^a)^{\alpha\beta}\partial_a \\
\{D_{\alpha i}, D^{\beta j'}\} &= 0,
\end{aligned} \tag{5.21}$$

where $\alpha = 1, \ldots, 4$ is a chiral spinor index and i, i' are doublet indices for the two SU(2)s. In this notation the field strength is $W_i{}^{i'}$ and the supercurrent is $J_{ij}^{i'j'} := Tr(W_{(i}{}^{i'}W_{j)}{}^{j'})$. The double-trace true $\frac{1}{2}$-BPS superfield is $B_{ijkl}^{i'j'k'l'} := J_{(ij}^{(i'j'}J_{kl)}^{k'l')}$. It obeys the constraint

$$D_{\alpha(i}B_{jklm)}^{j'k'l'm'} = 0, \tag{5.22}$$

together with a similar one for the upper indices. The one-half BPS Lagrangian 6-form starts at $L_{0,6}$. It is

$$L_{\alpha i\beta j\gamma k}^{\delta l'\varepsilon m'\phi n'} := \delta_{(\alpha}{}^{(\delta}\delta_b{}^{\varepsilon}B_{\gamma)ijk}^{\phi)l'm'n'}, \tag{5.23}$$

where

$$B_{\alpha ijk}^{\beta i'j'k'} := D_\alpha{}^l D_{l'}^\beta B_{ijkl}^{i'j'k'l'} . \tag{5.24}$$

There are two $Spin(1,5)$ representations here: a singlet and a 15, but it turns out that precisely this combination is required in order to obtain an element of $H_s^{0,6}$. Moreover, it is not difficult to show that this form cannot be shortened so that the cocycle for the true $\frac{1}{2}$-BPS invariant is different to that of the Lagrangian (which starts at $L_{1,5}$ in $D=6$).

As in the $D=7$ case, the double-trace $\frac{1}{4}$-BPS invariant can be constructed in terms of a pseudo-$\frac{1}{2}$-BPS superfield obeying (5.22). Again, it is formed by inserting a pair of contracted spacetime derivatives, one on each factor of J. We now have the task of testing for the cohomological triviality of the corresponding Lagrangian form, i.e., we need to try to write $L_{0,6} = d_1 K_{0,5} + t_0 K_{1,4}$. In contrast to the $D=7$ case, however, we find that one cannot do this.

The problem can be approached from different points of view. The first is to try to repeat what was done for $D=7$ by writing K in terms of bilinears of the supercurrent, but it turns out that no such $K_{0,5}$ and $K_{1,4}$ can be constructed in this way. Alternatively, we can observe that there are two true $\frac{1}{4}$-BPS bilinears that can be constructed from J,

$$\begin{aligned} C_{ijkl} &:= J_{(ij}^{i'j'} J_{kl)i'j'} \\ C'^{i'j'k'l'} &:= J_{(ij}^{(i'j'} J^{k'l')ij} . \end{aligned} \tag{5.25}$$

These superfields obey constraints of the type (5.22), C with respect to D and C' with respect to D'. There is another shortened bilinear that can be constructed from J; it is

$$S_{ij}^{i'j'} := J_{k(i}^{k'(i'} J_{j)l}^{j')l'} \varepsilon^{kl} \varepsilon_{k'l'} . \tag{5.26}$$

It obeys constraints that are of third order in D and D' separately. This is akin to the product of two supercurrents in $N=4, D=4$ which is protected as a superconformal field even though it is not BPS-shortened [28, 34].

The pseudo-$\frac{1}{2}$-BPS operator B can be written using four derivatives acting on any of these three superfields, up to a total spacetime derivative which is irrelevant under integration. One has

$$\begin{aligned} B_{ijkl}^{i'j'k'l'} &\sim D_{ijkl}^4 C'^{i'j'k'l'} \\ &\sim D^{4i'j'k'l'} C_{ijkl} \\ &\sim D_{\alpha\beta(ij}^2 D^{2\alpha\beta(i'j'} S_{kl)}^{k'l')} , \end{aligned} \tag{5.27}$$

where the D^4s are fourth-order in D and are totally symmetric on the internal indices, while the second-order D^2s are symmetric in the internal indices and anti-symmetric on the spinor indices. The $\frac{1}{4}$-BPS invariant can be written as a 12-θ integral of any of these, so that one might expect to be able to trivialize the cohomology by using any one of them in K. But this turns out to be not possible even if one includes all three at once.

We therefore conclude that the double-trace d^2F^4 invariant is actually protected in $N=2, D=6$ SYM even though a similar invariant is not protected in $D=7$. A key difference between the two cases is the larger R-symmetry group in $D=6$ which is more restrictive when it comes to constructing possible trivializing $(D-1)$-forms K.

This result is in agreement with the $D=6, L=3$ SYM calculations [9] for the double-trace divergences. This demonstrates the subtlety that non-renormalization theorems can display in allowing and disallowing apparently similar operators in different spacetime dimensions, and it gives hope for a similar understanding of the maximal supergravity $D=6 \leftrightarrow D=5$ $\frac{1}{8}$-BPS conundrum.

The evaluation of superspace cohomology groups for maximal supergravity is a more difficult problem, principally because of the larger R-symmetry groups, many of which have the disadvantage of being symplectic. There is also a conceptual issue to deal with arising from the on-shell vanishing of supergravity actions. Nevertheless, we can speculate as to the outcome of such investigations using maximal super Yang–Mills as a guide. It seems likely that similar cohomological arguments can be found to protect the $\frac{1}{8}$-BPS $(L=4,\ D=5)$ invariant. Do such arguments work in lower dimensions, invalidating the $\frac{1}{4}$-BPS $(L=4,\ D=5)$ and $\frac{1}{8}$-BPS $(L=6,$ $D=4)$ candidate counterterms? If so, this would mean that all BPS counterterms of any degree are protected in $N=8, D=4$ supergravity, leading to the anticipation that the first $D=4$ divergence could occur only as a full superspace integral, which would happen at $L=7$ loops. Such a counterterm was explicitly constructed in the linearized theory many years ago [35], although this is not manifestly invariant under the non-linear E_7 symmetry. Manifestly E_7 invariant candidates occur starting in $L=8$ loops [35, 47].

From the point of view of the non-renormalization theorems, it seems ultimately to be a losing game to hope for continued ultraviolet cancellations in maximal supergravity, in the face of the ever-rising naïve degree of divergence (5.1). That is, unless true “miracles” of cancellation start to turn up when the loop order finally gets high enough for supersymmetry to have run out of its remarkable infinity-suppressing properties. All one can say at the moment is that the jury is still out. Table 5.4 the current state of play for the known supergravity divergences.

Table 5.4. *Current (late 2009) status of known supergravity divergences*

Dimension D	11	10	8	7	6	5	4
Loop order L	2	2	1	2	3	≥ 6	≥ 5
General form	$\partial^{12}R^4$	$\partial^{10}R^4$	R^4	$\partial^4 R^4$	$\partial^6 R^4$	?	?

References

[1] L. Baulieu *et al.*, Finiteness properties of the $\mathcal{N}=4$ super-Yang–Mills theory in supersymmetric gauge, *Nucl. Phys.* **B753**, 252–72 (2006).

[2] L. Baulieu *et al.*, Shadow fields and local supersymmetric gauges, *Nucl. Phys.* **B753**, 273–94 (2006).

[3] L. Baulieu *et al.*, Ten-dimensional super-Yang–Mills with nine off-shell supersymmetries, *Phys. Lett.* **B658**, 249–52 (2008).

[4] Z. Bern *et al.*, One-loop n-point gauge theory amplitudes, unitarity and collinear limits, *Nucl. Phys.* **B425**, 217–60 (1994).

[5] Z. Bern *et al.*, On the relationship between Yang–Mills theory and gravity and its implication for ultraviolet divergences, *Nucl. Phys.* **B530**, 401–56 (1998).

[6] Z. Bern *et al.*, The four-loop planar amplitude and cusp anomalous dimension in maximally supersymmetric Yang–Mills theory, *Phys. Rev.* **D75**, 085010 (2007).

[7] Z. Bern *et al.*, Three-loop superfiniteness of $\mathcal{N}=8$ supergravity, *Phys. Rev. Lett.* **98**, 161303 (2007).

[8] Z. Bern *et al.*, Manifest ultraviolet behavior for the three-loop four-point amplitude of $\mathcal{N}=8$ supergravity, *Phys. Rev.* **D78**, 105019 (2008).

[9] Z. Bern *et al.*, Multi-loop amplitudes with maximal supersymmetry, Talk by L. J. Dixon at the International Workshop on Gauge and String Amplitudes, IPPP Durham, 30th March to 3rd April 2009. (Available on-line at http://conference.ippp.dur.ac.uk) (2009).

[10] Z. Bern *et al.*, The ultraviolet behavior of $\mathcal{N}=8$ supergravity at four loops, *Phys. Rev. Lett.* **103**, 081301 (2009).

[11] N. Berkovits, ICTP lectures on covariant quantization of the superstring, hep-th/0209059 (2002).

[12] N. Berkovits, New higher-derivative R^4 theorems, *Phys. Rev. Lett.* **98**, 211601 (2007).

[13] N. Berkovits and P. S. Howe, The cohomology of superspace, pure spinors and invariant integrals, arxiv:0803.3024 (2008).

[14] N. Berkovits *et al.*, Non-renormalization conditions for four-gluon scattering in supersymmetric string and field theory, *JHEP* **11**, 063 (2009).

[15] L. Bonara, P. Pasti and M. Tonin, Superspace formulation of 10-D SUGRA+SYM theory a la Green–Schwarz, *Phys. Lett.* **B188**, 335 (1987).

[16] L. Bonara, P. Pasti and M. Tonin, Chiral anomalies in higher dimensional supersymmetric theories, *Nucl. Phys.* **B286**, 150 (1987).

[17] G. Bossard, P. S. Howe and K. S. Stelle, The ultra-violet question in maximally supersymmetric field theories, *Gen. Rel. Grav.* **41**, 919–81 (2009).

[18] G. Bossard, P. S. Howe and K. S. Stelle, A note on the UV behaviour of maximally supersymmetric Yang–Mills theories, *Phys. Lett.* **B682**, 137–42 (2009).

[19] L. Brink, O. Lindgren and B. E. W. Nilsson, The ultraviolet finiteness of the $\mathcal{N}=4$ Yang–Mills theory, *Phys. Lett.* **B123**, 323 (1983).
[20] M. Cederwall, B.E.W. Nilsson and D. Tsimpis, Spinorial cohomology and maximally supersymmetric theories, *JHEP* **02**, 009 (2002).
[21] E. Cremmer *et al.*, Spontaneous symmetry breaking and Higgs effect in supergravity without cosmological constant, *Nucl. Phys.* **B147**, 105 (1979).
[22] F. Delduc and J. McCabe, The quantization of $\mathcal{N}=3$ super Yang–Mills off-shell in harmonic superspace, *Class. Quant. Grav.* **06**, 233 (1989).
[23] S. Deser and B. Zumino, Consistent supergravity, *Phys. Lett.* **B62**, 335 (1976).
[24] S. Deser, Renormalizability properties of supergravity, *Phys. Rev. Lett.* **38**, 527 (1977).
[25] J. A. Dixon, Supersymmetry is full of holes, *Class. Quant. Grav.* **07**, 1511–21 (1990).
[26] J. M. Drummond *et al.*, Integral invariants in $\mathcal{N}=4$ SYM and the effective action for coincident D-branes, *JHEP* **08**, 016 (2003).
[27] M. J. Duff and D. J. Toms, Kaluza–Klein kounterterms. Presented at the 2nd Europhysics Study Conference on Unification of Fundamental Interactions, Erice, Sicily, Oct 6–14, 1981.
[28] B. Eden *et al.*, Partial non-renormalisation of the stress-tensor four-point function in $\mathcal{N}=4$ SYM and AdS/CFT, *Nucl. Phys.* **B607**, 191–212 (2001).
[29] D. Z. Freedman, P. van Nieuwenhuizen and S. Ferrara, Progress toward a theory of supergravity, *Phys. Rev.*, **D13**, 3214–18 (1976).
[30] A. Galperin *et al.*, Unconstrained off-shell $N=3$ supersymmetric Yang–Mills theory, *Class. Quant. Grav.* **02**, 155 (1985).
[31] A. Galperin *et al.*, $\mathcal{N}=3$ supersymmetric gauge theory, *Phys. Lett.* **B151**, 215–18 (1985).
[32] S. J. Gates Jr. *et al.*, Component actions from curved superspace: normal coordinates and ectoplasm, *Phys. Lett.* **B421**, 203–10 (1998).
[33] M. B. Green, J. G. Russo and P. Vanhove, Ultraviolet properties of maximal supergravity, *Phys. Rev. Lett.* **98**, 131602 (2007).
[34] P. J. Heslop and P. S. Howe, A note on composite operators in $\mathcal{N}=4$ SYM, *Phys. Lett.* **B516**, 367–75 (2001).
[35] P. S. Howe and U. Lindstrom, Higher order invariants in extended supergravity, *Nucl. Phys.* **B181**, 487 (1981).
[36] P. S. Howe, K. S. Stelle and P. K. Townsend, Superactions, *Nucl. Phys.* **B191**, 445 (1981).
[37] P. S. Howe, K. S. Stelle and P. K. Townsend, The relaxed hypermultiplet: an unconstrained $\mathcal{N}=2$ superfield theory, *Nucl. Phys.* **B214**, 519 (1983).
[38] P. S. Howe, K. S. Stelle and P. K. Townsend, Miraculous ultraviolet cancellations in supersymmetry made manifest, *Nucl. Phys.* **B236**, 125 (1984).
[39] P. S. Howe, G. Papadopoulos and K.S. Stelle, The background field method and the nonlinear sigma model, *Nucl. Phys.* **B296**, 26 (1988).
[40] P. S. Howe and K. S. Stelle, The ultraviolet properties of supersymmetric field theories, *Int. J. Mod. Phys.* **A4**, 1871 (1989).
[41] P. S. Howe, U. Lindstrom and P. White, Anomalies and renormalization in the BRST-BV framework, *Phys. Lett.* **B246**, 430–34 (1990).
[42] P. S. Howe, Pure spinors, function superspaces and supergravity theories in ten-dimensions and eleven-dimensions, *Phys. Lett.* **B273**, 90–94 (1991).
[43] P. S. Howe, Pure spinors lines in superspace and ten-dimensional supersymmetric theories, *Phys. Lett.* **B258**, 141–4 (1991).
[44] P. S. Howe, O. Raetzel and E. Sezgin, On brane actions and superembeddings, *JHEP* **08**, 011 (1998).

[45] P. S. Howe and K. S. Stelle, Supersymmetry counterterms revisited, *Phys. Lett.* **B554**, 190–96 (2003).
[46] P. S. Howe and D. Tsimpis, On higher-order corrections in M theory, *JHEP* **09**, 038 (2003).
[47] R. E. Kallosh, Counterterms in extended supergravities, *Phys. Lett.* **B99**, 122–7 (1981).
[48] R. Kallosh, On a possibility of a UV finite $\mathcal{N} = 8$ supergravity, arxiv:0808.2310 (2008).
[49] S. Mandelstam, Light cone superspace and the ultraviolet finiteness of the $\mathcal{N} = 4$ model, *Nucl. Phys.* **B213**, 149–68 (1983).
[50] N. Marcus and A. Sagnotti, The ultraviolet behavior of $\mathcal{N} = 4$ Yang–Mills and the power counting of extended superspace, *Nucl. Phys.* **B256**, 77 (1985).
[51] G. 't Hooft and M. J. G. Veltman, One loop divergencies in the theory of gravitation, *Ann. Poincare Phys. Theor.* **A20**, 69–94 (1974).

6

Cosmological quantum billiards

AXEL KLEINSCHMIDT & HERMANN NICOLAI

The minisuperspace quantization of $D = 11$ supergravity is equivalent to the quantization of an $E_{10}/K(E_{10})$ coset space sigma model, when the latter is restricted to the E_{10} Cartan subalgebra. As a consequence, the wavefunctions solving the relevant minisuperspace Wheeler–DeWitt equation involve automorphic (Maass wave) forms under the modular group $W^{+}(E_{10}) \cong PSL_2(\mathtt{O})$. Using Dirichlet boundary conditions on the billiard domain a general inequality for the Laplace eigenvalues of these automorphic forms is derived, entailing a wave function of the universe that is generically complex and always tends to zero when approaching the initial singularity. The significance of these properties for the nature of singularities in quantum cosmology in comparison with other approaches is discussed. The present approach also offers interesting new perspectives on some longstanding issues in canonical quantum gravity.

6.1 Introduction

The present chapter is based on [1], and elaborates on several issues and arguments that were not fully spelled out there. In that work, a first step was taken towards quantization of the one-dimensional "geodesic" $E_{10}/K(E_{10})$ coset model which had been proposed in [2] as a model of M-theory. Here, E_{10} denotes the hyperbolic Kac–Moody group E_{10} which is an infinite-dimensional extension of the exceptional Lie group E_8, and plays a similarly distinguished role among the infinite-dimensional Lie algebras as E_8 does among the finite-dimensional ones.

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press. © Cambridge University Press 2012.

The proposal of [2] had its roots both in the appearance of so-called "hidden symmetries" of exceptional type in the dimensional reduction of maximal supergravity to lower dimensions [3], as well as in the celebrated analysis of Belinskii, Khalatnikov and Lifshitz (BKL) [4] of the gravitational field equations in the vicinity of a generic space-like (cosmological) singularity. According to the basic hypothesis underlying this analysis the causal decoupling of spatial points near the space-like singularity[1] effectively leads to a dimensional reduction whereby the equations of motion become ultralocal in space, and the dynamics should therefore be describable in terms of a (continuous) superposition of one-dimensional systems, one for each spatial point. More specifically, in this approximation the dynamics at each spatial point can be described by a sequence of Kasner regimes, such that in the strict limit towards the singularity, the Kasner behavior is interspersed with hard reflections of the logarithms of the spatial scale factors off infinite potential walls [5,6]. This generic behavior has been termed "cosmological billiards." The geometry of the billiard table and the configuration of the walls ("cushions") of the billiard table are determined by the dimension and the matter content of the theory [7,8]. Likewise the occurrence or non-occurrence of chaotic oscillations near the singularity depends on this configuration. In particular, for $D = 11$ supergravity it was shown in [9] that the billiard domain is the fundamental Weyl chamber $\mathcal{C}$ of the "maximally extended" hyperbolic Kac–Moody group E_{10}. The volume of this fundamental Weyl chamber is finite, implying chaotic behavior [9,10]. The emergence of the hyperbolic Kac–Moody algebra E_{10} in this context is also in line with its conjectured appearance in the dimensional reduction of $D = 11$ supergravity to one time dimension [11].

Reference [2] goes beyond the standard BKL analysis, as well as the original conjecture [11], in that it establishes a correspondence at the classical level between a truncated gradient expansion of the $D = 11$ supergravity equations of motion near the space-like singularity and an expansion in heights of roots of a one-dimensional constrained "geodesic" $E_{10}/K(E_{10})$ coset space model. The cosmological billiards approximation then corresponds to the restriction of this coset model to the Cartan subalgebra of E_{10}. Going beyond this billiard approximation involves bringing in spatial dependence, in such a way that a "small tension" expansion in spatial gradients *à la* BKL gets converted into a level expansion of the E_{10} Lie algebra. However, the correspondence between the field equations on the one hand, and the $E_{10}/K(E_{10})$ model on the other hand, codified in a "dictionary," has so far only been shown to work up to first-order spatial gradients. The proper inclusion of higher-order spatial gradients, and thus the emergence of a space-time field theory

[1] This is the same decoupling that later came to be associated with the so-called "horizon problem" of inflationary cosmology.

from a "pregeometrical" scheme, remains an outstanding problem, in spite of the fact that the E_{10} Lie algebra contains all the "gradient representations" that would be needed for a Taylor expansion around a given spatial point [2].

Quantizing M-theory in the E_{10} framework thus amounts to setting up and solving a Wheeler–DeWitt equation for the full $E_{10}/K(E_{10})$ model, and imposing the subsidiary constraints corresponding to the canonical (diffeomorphism, Gauss, etc.) constraints of the usual canonical approach. As a very first step, [1] solves the quantum constraints for $D=11$ supergravity for the 10 spatial scale factors, and for their fermionic "superpartners," in compliance with the supersymmetry constraint. The resulting *quantum cosmological billiard* is a variant of the "minisuperspace" quantization of gravity pioneered in [12–14] and further developed in many works, see in particular [15–20] and references therein.[2] The essential new ingredient in the present construction is the *arithmetic structure* provided by E_{10} and its Weyl group, whose relevance in the context of Einstein gravity was pointed out and explored in [22]; see also [23] for an ansatz based on superconformal quantum mechanics, which has some similarities to the present approach. We believe that the key issues with the proposal of [2], in particular the conjectured emergence of classical space-time out of a pregeometrical quantum gravity phase, and the question of how the "dictionary" of [2] can be extended to the full E_{10} algebra, cannot be resolved within the framework of classical field equations, but will require quantization of the $E_{10}/K(E_{10})$ coset model. Equally important, the "resolution" of the cosmological singularity, a key issue of modern research in canonical quantum gravity, is expected to involve quantum theory in an essential way. In addition, it will almost certainly require new concepts besides the transition to the quantum theory, with the question of what happens to classical space-time near the singularity as the core problem.

The main results reported in [1] and in the present chapter are:

- The quantum cosmological billiard problem is well-posed and a Hilbert space (with positive definite metric) can be defined.
- The solutions to the bosonic Wheeler–DeWitt (WDW) equation of $D=11$ supergravity can be described as odd Maass wave forms of the "modular group" $W^+(E_{10}) \cong PSL(2, 0)$.
- With the appropriate Dirichlet boundary conditions all solutions of the WDW equation, a.k.a. "wavefunctions of the universe," can be shown generally to vanish rapidly near a space-like singularity while remaining complex and oscillating.

[2] We remind readers that the "super" in "minisuperspace" refers to Wheeler's moduli space of 3-geometries and is a purely bosonic concept. It therefore has no relation to the notion of "supersymmetry" relating bosons and fermions, which is also discussed in this chapter.

- The analysis can be extended to include the fermionic degrees of freedom without changing the conclusions.

Possible extensions and the eventual significance of these results for quantum cosmology are discussed at the end of this chapter.

6.2 Minisuperspace quantization

We first consider the bosonic variables. We proceed from the following metric ansatz, as appropriate for a cosmological billiard for a $(d+1)$-dimensional space-time

$$ds^2 = -N^2 dt^2 + \sum_{a=1}^{d} e^{-2\beta^a} \theta^a \otimes \theta^a \,, \tag{6.1}$$

where we keep the spatial dimension d arbitrary (but always $d \geq 3$) for the moment, and will specialize to $D = 11$ supergravity (and thus $d = 10$) only later. $\theta^a \equiv \mathcal{N}^a{}_m dx^m$ is a spatial frame in an Iwasawa decomposition of the metric, as explained in [7]. Substituting the above ansatz into the Einstein action, one arrives at the kinetic term

$$\mathcal{L}_{\text{kin}} = \frac{1}{2} n^{-1} \sum_{a,b=1}^{d} \dot{\beta}^a G_{ab} \dot{\beta}^b \tag{6.2}$$

in terms of the new lapse $n \equiv N/\sqrt{g}$ (the spatial volume is $\sqrt{g} = \exp[-\sum_a \beta^a]$) and the Lorentzian DeWitt metric

$$\dot{\beta}^a G_{ab} \dot{\beta}^b \equiv \sum_{a=1}^{d} (\dot{\beta}^a)^2 - \left(\sum_{a=1}^{d} \dot{\beta}^a \right)^2 . \tag{6.3}$$

As is well known, the DeWitt metric, with the factor (-1) in front of the second term, is distinguished by several uniqueness properties that are discussed in [24]. Here, it will be essential that for $d = 10$ this metric coincides with the restriction of the Cartan–Killing metric of E_{10} to its Cartan subalgebra. It can now be shown [7] that the remaining contributions to the Hamiltonian constraint at a given spatial point can be combined into an "effective potential" of the generic form

$$V_{\text{eff}} = \sum_A c_A(Q, P, \partial\beta, \partial Q) \exp\big(-2w_A(\beta)\big) \tag{6.4}$$

where (Q, P) are the (canonical) variables corresponding to all degrees of freedom other than the scale factors β^a, the ∂ stands for *spatial* gradients, and w_A in the exponent are linear forms, called *wall forms*,

$$w_A(\beta) \equiv G_{ab} w_A^a \beta^b \tag{6.5}$$

with the DeWitt metric G_{ab} introduced above. In the limit towards the singularity $\beta \to \infty$, the exponential walls become "sharp," and the dynamics is dominated by a set of "nearest" walls. These make up the "cushions" of a billiard table, and can be viewed as the result of "integrating out" the off-diagonal metric and the matter degrees of freedom, as explained in [7]. In the strict limit towards the singularity they are simply given by time-like hyperplanes in the forward lightcone in β-space, which are determined by the linear equations $w_A(\beta) = 0$. The spatial ultralocality of the BKL limit thus reduces the gravitational model to a classical mechanics system of a relativistic billiard ball described by the β^a variables moving on straight null lines in the Lorentzian space with metric G_{ab} until hitting a billiard table wall corresponding to a hyperplane. The straight-line segments of the billiard motion are the Kasner regimes, while the reflections are usually referred to as "Kasner bounces." We repeat that there is one such system for each spatial point $\mathbf{x}$, and these systems are all decoupled.

The canonical bosonic variables of the billiard system are β^a and their canonically conjugate momenta π_a, viz.

$$\pi_a := \frac{\partial \mathcal{L}}{\partial \dot{\beta}^a} = G_{ab} \dot{\beta}^b \tag{6.6}$$

where we set $n = 1$ from now on.[3] The Hamiltonian is

$$\mathcal{H}_0 = \frac{1}{2} \pi_a G^{ab} \pi_b \tag{6.7}$$

with the inverse DeWitt metric G^{ab}. The effective potential (6.4) has disappeared as we have taken the BKL limit. Before quantization, we perform the following change of variables by means of which the billiard motion is projected onto the unit hyperboloid in β-space [7]:

$$\beta^a = \rho \omega^a \,, \quad \omega^a G_{ab} \omega^b = -1 \,, \quad \rho^2 = -\beta^a G_{ab} \beta^b \,, \tag{6.8}$$

where ρ is the "radial" direction in the future lightcone and $\omega^a = \omega^a(z)$ are expressible as functions of $d - 1$ coordinates z on the unit hyperboloid. The limit towards the singularity is $\rho \to \infty$ in these variables. The Wheeler–DeWitt operator thus

[3] With this choice of gauge, t becomes a "Zeno-like" time coordinate, for which the singularity is at $t = +\infty$. This time is related to physical (proper) time T by $t \sim -\log T$.

takes the form

$$\mathcal{H}_0 \equiv G^{ab}\partial_a\partial_b = -\rho^{1-d}\frac{\partial}{\partial\rho}\left(\rho^{d-1}\frac{\partial}{\partial\rho}\right) + \rho^{-2}\Delta_{\text{LB}}\,, \tag{6.9}$$

where Δ_{LB} is the Laplace–Beltrami operator on the $(d-1)$-dimensional unit hyperboloid. We emphasize that ordering ambiguities are entirely absent in this expression, as is manifest in the β^a variables in terms of which the WDW equation is just the free Klein–Gordon equation. The same holds true for the variables $(\rho, \omega^a(z))$ because the expression in the new coordinates is unambiguously determined by the coordinate transformation (6.8) (as it would be in any other coordinate system).

The minisuperspace WDW equation therefore reads

$$\mathcal{H}_0\Phi(\rho, z) = 0 \tag{6.10}$$

for the "wavefunction of the universe" $\Phi(\rho, z)$. As usual (see e.g. [21]) one can adopt ρ as a time coordinate in the initially "timeless" WDW equation, with the standard (Klein–Gordon-like) invariant inner product

$$(\Phi_1, \Phi_2) = i\int d\Sigma^a\,\Phi_1^*\overset{\leftrightarrow}{\partial_a}\Phi_2 \tag{6.11}$$

where the integral is to be taken over a space-like hypersurface inside the forward lightcone in β-space. "Invariance" means that this scalar product does not depend on the shape of the space-like hypersurface, so we can for instance choose any of the hyperboloids $\rho = \text{const}$.

In order to construct solutions we separate variables by means of the ansatz $\Phi(\rho, z) = R(\rho)F(z)$ [14, 16]. For any eigenfunction $F(z)$ obeying

$$-\Delta_{\text{LB}}F(z) = EF(z) \tag{6.12}$$

the associated radial equation is solved by

$$R_\pm(\rho) = \rho^{-\frac{d-2}{2}}e^{\pm i\sqrt{E-\left(\frac{d-2}{2}\right)^2}\log\rho}\,. \tag{6.13}$$

Positive-frequency waves emanating from the singularity correspond to $R_-(\rho)$ and have positive inner product (6.11). It is important here that one can consistently restrict to positive norm wavefunctions: the potential which might scatter an initially positive norm state into a negative norm state is here effectively replaced by a set of boundary conditions on the wavefunction, and hence there is no "Klein paradox." As we will see, the same feature continues to hold for the full E_{10} WDW operator, in marked contrast to the standard WDW operator.

To study the eigenvalues of the Laplace–Beltrami operator on the unit hyperboloid we use a generalized upper half plane model $z = (\vec{u}, v)$ for the unit hyperboloid with coordinates $\vec{u} \in \mathbb{R}^{d-2}$ and $v \in \mathbb{R}_{>0}$. The relevant coordinate transformation is obtained by first diagonalizing the DeWitt metric (6.3) in terms of Minkowskian coordinates $\widetilde{\beta}^a$, such that

$$G_{ab}\beta^a\beta^b \equiv -\widetilde{\beta}^+\widetilde{\beta}^- + \sum_{j=1}^{d-2}\widetilde{\beta}^j\widetilde{\beta}^j\,, \tag{6.14}$$

where we have used the last two directions for forming lightcone coordinates. Then the forward unit hyperboloid (with $\widetilde{\beta}^{\pm} > 0$) is coordinatized by

$$\widetilde{\beta}^+ = \frac{1}{v}\,, \quad \widetilde{\beta}^- = v + \frac{\vec{u}^{\,2}}{v}\,, \quad \widetilde{\beta}^j = \frac{u^j}{v} \qquad (v > 0). \tag{6.15}$$

The metric induced on the unit hyperboloid is easily calculated to be the Poincaré metric on the generalized upper half plane

$$ds^2 = \frac{dv^2 + d\vec{u}^{\,2}}{v^2} \quad \Rightarrow \quad d\mathrm{vol}(z) \equiv v^{1-d}\,dv d^{d-2}u \tag{6.16}$$

such that the Laplace–Beltrami operator becomes

$$\Delta_{\mathrm{LB}} = v^{d-1}\partial_v\left(v^{3-d}\partial_v\right) + v^2\partial_{\vec{u}}^2\,. \tag{6.17}$$

For the spectral problem we must specify boundary conditions. For the cosmological billiard, these are provided by infinite (“sharp”) potential walls which encapsulate the effect of spatial inhomogeneities and matter fields near the spacelike singularity, as explained above (see [7, 8] for details). Following the original suggestion of [14], we are thus led to impose the vanishing of the wavefunction on the boundary of the fundamental domain specified by these walls. Accordingly, let $F(z)$ be any function on the hyperboloid satisfying (6.12) with *Dirichlet conditions* at the boundaries of this domain (in contrast to [16, 22], where Neumann boundary conditions are assumed). A direct generalization of the arguments on p. 28 of Ref. [25] gives

$$-(\Delta_{\mathrm{LB}}F, F) \geq \int dv\, d^{d-2}u\, v^{3-d}(\partial_v F)^2 \tag{6.18}$$

with (6.12) and (6.17). Considering also

$$\begin{aligned}(F, F) &\equiv \int d\mathrm{vol}(z)F^2(z) = \int dv\, d^{d-2}u\, v^{1-d}F^2 \\ &= \frac{2}{d-2}\int dv\, d^{d-2}u\, v^{2-d}F\partial_v F\,,\end{aligned} \tag{6.19}$$

the use of the Cauchy–Schwarz inequality entails

$$E \geq \left(\frac{d-2}{2}\right)^2 . \tag{6.20}$$

From the explicit solution (6.13) we thus conclude that $R_{\pm}(\rho) \to 0$ when $\rho \to \infty$, and therefore *the full wavefunction and all its ρ derivatives tend to zero near the singularity*. Evidently, this result hinges on the peculiar form of the differential operator in the (ρ, z) variables in (6.9), which itself is uniquely determined by the form of the operator in β-coordinates. It would not be valid if we were allowed to move around the ρ factors in the differential operator of (6.9).

While the wavefunction would also vanish for Neumann boundary conditions (for which $E \geq 0$) with the given ordering, the inequality (6.20) furthermore ensures that the full wavefunction is generically complex and oscillating. Let us point out here that this result may be of relevance to a long-standing issue in canonical gravity, namely the question of why and how the *real* WDW equation should give rise to *complex* wavefunctions [26–28]. As explained there, the complexity of the wavefunction is intimately linked to the emergence of a *directed* time in canonical gravity. More specifically, admitting only positive-norm wavefunctions corresponds to choosing an "arrow of time" (real wavefunctions have vanishing norm with the product (6.11), and would thus not select a time direction). Let us repeat that restricting to positive norm states would be inconsistent for the standard WDW equation with a potential even in minisuperspace quantization. Here, the potential has effectively disappeared in the BKL limit, leaving its trace only via the boundary conditions, so the restriction is consistent.

6.3 Automorphy and the E_{10} Weyl group

Whereas the discussion above was valid for gravity in any space-time of dimension $d+1$, we now focus on maximal supergravity in 11 space-time dimensions. For the bosonic sector of maximal supergravity, the wavefunctions can be further analyzed by exploiting the underlying symmetry encoded in the Weyl group $W(E_{10})$ and its arithmetic properties, and in particular the new links between hyperbolic Weyl groups and generalized modular groups uncovered in [29]. The Weyl reflections that the classical particle is subjected to when colliding with one of the walls are norm preserving, and therefore the reflections can be projected to any hyperboloid of constant ρ, inducing a non-linear action on the coordinates z (given in (6.23) below for the fundamental reflections). For physical amplitudes to be invariant under the Weyl group, the full wavefunction must transform as follows:

$$\Phi(\beta) = \pm\Phi(w_I(\beta)) \quad \Leftrightarrow \quad \Phi(\rho, z) = \pm\Phi\big(\rho, w_I(z)\big) \tag{6.21}$$

for the 10 generating fundamental reflections w_I of $W(E_{10})$, labeled by $I = -1, 0, 1, \dots, 8$. Restricting the wavefunction to the fundamental Weyl chamber, one easily checks that the plus sign in (6.21) corresponds to Neumann boundary conditions, and the minus sign to Dirichlet conditions (which we adopt here). From (6.21) it follows that $\Phi(\rho, z)$ is invariant under *even* Weyl transformations $s \in W^+(E_{10})$ irrespective of the chosen boundary conditions.

Choosing coordinates as in (6.15), the relevant variables now live in a nine-dimensional "octonionic upper half plane" with coordinate

$$z = u + \mathrm{i}v, \qquad u \equiv \vec{u} \in \mathbb{O}, \tag{6.22}$$

where $\mathbb{O}$ is the non-commutative and non-associative algebra of *octonions*, while v is still real and positive. Next, we recall [29–31] that the 240 roots of E_8 can be represented by unit octonions; more precisely, these are the 240 *units* in the non-commutative and non-associative ring of *integral octonions* $\mathtt{O}$ called *"octavians,"* see [31]. Denoting by ε_j (for $j = 1, \dots, 8$) the eight simple roots and by θ the highest root of E_8, respectively, expressed as unit octonions, we arrive at the following *modular realization* of the E_{10} Weyl transformations on the nine-dimensional unit hyperboloid: the 10 fundamental reflections of $W(E_{10})$ act as

$$w_{-1}(z) = \frac{1}{\bar{z}}, \quad w_0(z) = -\theta\bar{z}\theta + \theta, \quad w_j(z) = -\varepsilon_j \bar{z} \varepsilon_j \tag{6.23}$$

where $\bar{z} := \bar{u} - \mathrm{i}v$, with $\mathrm{i}u = \bar{u}\mathrm{i}$ in accordance with Cayley–Dickson doubling [31].[4] Observe that, despite the non-associativity of the octonions, there is no need to put parentheses in (6.23) by virtue of the alternativity of the octonions.

In the present context, the formulas (6.23) represent the most general (and most sophisticated!) modular realization of a Weyl group, but there are corresponding versions for the other division algebras, with the quaternions $\mathbb{H}$ for $d = 6$, and the complex numbers $\mathbb{C}$ for $d = 4$ (with corresponding "integers," see [29] for details). The simplest case is $\mathbb{A} = \mathbb{R}$, which corresponds to pure gravity in four space-time dimensions ($d = 3$). In this case $u \in \mathbb{R}$, and the formulas (6.23) reduce to those familiar from complex analysis, namely

$$z \mapsto \frac{1}{\bar{z}}, \quad z \mapsto -\bar{z} + 1, \quad z \mapsto -\bar{z}, \tag{6.24}$$

[4] We recall that the Dickson doubling of a normed algebra $\mathbb{A}$ with conjugation associates to doubled elements $a + \mathrm{i}b, c + \mathrm{i}d \in \mathbb{A} + \mathrm{i}\mathbb{A}$ the product (see [31])

$$(a + \mathrm{i}b)(c + \mathrm{i}d) = (ac - \bar{d}b) + \mathrm{i}(cb + \bar{a}d).$$

The Hurwitz theorem states that, starting from the real numbers $\mathbb{R}$, this process generates the division algebras $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ of the real, complex, quaternionic and octonionic numbers, respectively. Further application generates algebras with zero divisors.

generating the group $PGL_2(\mathbb{Z})$. For *even* Weyl transformations, we reobtain the standard modular group $PSL_2(\mathbb{Z})$ generated by

$$S(z) \equiv (w_{-1}w_1)(z) = -1/z\,, \qquad T(z) \equiv (w_0 w_1)(z) = z+1\,. \tag{6.25}$$

Consequently, for pure gravity in four space-time dimensions, the relevant eigenfunctions of the minisuperspace WDW operator are automorphic forms with respect to the standard modular group $PSL_2(\mathbb{Z})$, as already pointed out in [22].

For the maximally supersymmetric theory, on the other hand, the even Weyl group $W^+(E_{10})$ is isomorphic to the "modular group" $PSL_2(\mathtt{O})$ over the octavians, where $PSL_2(\mathtt{O})$ is *defined* by iterating the action of (6.23) an even number of times [29]. Accordingly, for maximal supergravity the bosonic wavefunctions $\Phi(\rho,z)$ are *odd Maass wave forms* for $PSL_2(\mathtt{O})$, that is, invariant eigenfunctions of the Laplace–Beltrami operator transforming with a minus sign in (6.21) under the extension $W(E_{10})$ of $PSL_2(\mathtt{O})$. Properly understanding the "modular group" $PSL_2(\mathtt{O})$ and the associated modular functions remains an outstanding mathematical challenge, see [32, 33] for an introduction (and [25] for the $PSL_2(\mathbb{Z})$ theory). For the groups $PSL_2(\mathbb{Z})$ and $PSL_2(\mathbb{Z}[i])$ the (purely discrete) spectra of odd Maass wave forms have been investigated numerically in [34–38].

One important feature of (6.23) should be emphasized: supplementing the seven imaginary units of $\mathbb{O}$ by another imaginary unit i, it would appear that we have to enlarge the octonions to *sedenions*, a system of hypercomplex numbers with 15 imaginary units, which by Hurwitz's theorem is no longer a division algebra (that is, has zero divisors). Remarkably, however, the formulas (6.23) are such that with the iterated action of (6.23) we never need to introduce any further imaginary units beyond i and the seven octonionic ones. In other words, the transformations (6.23) do not move z out of the nine-dimensional generalized upper half plane. In particular, the doubling rule ensures that $\bar{z}z = v^2 + |u|^2 \in \mathbb{R}_+$ so that the inverse $1/\bar{z}$ also stays in this plane.

By modular invariance, the wavefunctions can be restricted to the fundamental domain of the action of $W(E_{10})$ and, conversely, their modular property defines them on the whole hyperboloid. The Klein–Gordon inner product (6.11) must likewise be restricted to the fundamental chamber

$$(\Phi_1, \Phi_2) = i \int_{\mathcal{F}} d\mathrm{vol}(z)\rho^{d-1}\Phi_1^* \overset{\leftrightarrow}{\partial}_\rho \Phi_2, \tag{6.26}$$

where $\mathcal{F}$ is the intersection of $\mathcal{C}$ with the unit hyperboloid; accordingly, the "cushions" of the billiard are obtained by intersecting the hyperplanes $w_A(\beta) = 0$ with the unit hyperboloid, such that for pure Einstein gravity ($d = 3$) one ends up with the projected billiard table shown in Figure 6.1. The restriction of the scalar product to the fundamental domain is necessary, as the integral over the whole hyperboloid

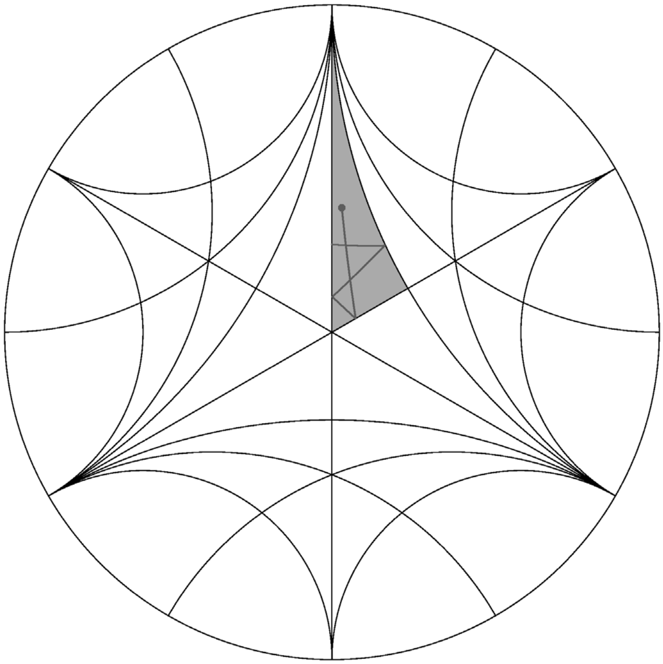


Figure 6.1 The projected billiard table of standard general relativity (in gray) on two-dimensional hyperbolic space. There are infinitely many equivalent tables related by Weyl reflections.

would be infinite for functions obeying (6.21), and the product would be ill-defined. This infinity is analogous to the one that arises in the calculation of the one-loop amplitude in string theory, and there as well, the product is rendered finite upon "division" by the modular group $PSL_2(\mathbb{Z})$. We thus have at hand an analog of this mechanism in canonical quantum gravity. We note that modular invariance is a distinctive feature of string theory not shared by the quantum field theory of point-like particles and arguably the "real" reason behind the conjectured finiteness of string theory.

6.4 Classical and quantum chaos

The fundamental region defined by the billiard walls as a subset of the unit hyperboloid (a hyperbolic space of constant negative curvature) has finite volume

$$\mathrm{Vol}(\mathcal{F}) = \int_{\mathcal{F}} d\mathrm{vol}(z) < \infty \tag{6.27}$$

in spite of the fact that the domain $\mathcal{F}$ extends to infinity (has a "cusp"). The finiteness of the fundamental domain is a consequence of the hyperbolicity of the Kac–Moody algeba E_{10} (whereas fundamental domains for non-hyperbolic indefinite Kac–Moody algebras have infinite volume). Such finite-volume billiards have been known and studied for a long time, as they are known to exhibit classical chaos. They are also the prototypical examples for studying the transition from classical to quantum chaos, and especially the question of how the presence of classical chaos is reflected in the spectra of the corresponding (unitary) Hamiltonian operators, see for example [39–42]. One of the remarkable results of these investigations is that there is a qualitative difference between the wavefunctions of classically ergodic and classically periodic orbits: the latter have very drastic (density) fluctuations whereas the former appear more like randomized Gaussians [39] and can be called quantum ergodic. Another notable feature is the appearance of so-called "scars" as remnants of classically periodic orbits [41, 43].[5] These are regions of (relatively) high probability in position space which appear related to the positions of classically periodic orbits.

However, there are two main differences between these studies and the cosmological (quantum) billiards considered here, viz.

- The cosmological billiard is *relativistic*, that is, the classical evolution follows a Klein–Gordon-like equation, instead of a non-relativistic Schrödinger equation (but see [44] for a discussion of relativistic neutrino billiards).
- In the β-space description, the walls of the billiard move away from one another, and one would thus have to solve the equation with *time-dependent* boundary conditions in the β variables. By contrast, the projection (6.8) allows us to reformulate the problem with static boundary conditions, at the expense of modifying the ρ-dependent part of the relativistic Hamiltonian $-\partial_\rho^2$ to the right-hand side of (6.9). For pure Einstein gravity in four dimensions, the resulting fixed walls billiard system is displayed in Figure 6.1.

Within this new setting it would be of much interest to study the fate of a generic wavepacket in our cosmological billiard system. The expectation is that the quantum theory "washes out" the classical chaos in the sense that any initially localized wavepacket will eventually disperse when approaching the singularity, such that the asymptotic (for $\rho \to \infty$) wavefunction will be spread evenly over $\mathcal{F}$. For the non-relativistic case some studies of the evolution of wavepackets can be found in [45], where the focus was on non-generic classically periodic configurations.

The particular case of interest in M-theory cosmology possesses a nine-dimensional fundamental domain and has apparently not been considered in the

[5] See [22] for a possible physical interpretation of these "scars" in quantum cosmology.

literature so far. The chaotic quantum billiard being merely the quantum theory of a *finite-dimensional* subsystem, corresponding to the Cartan subalgebra of an infinite-dimensional Kac–Moody algebra, such a study would however represent only a first step towards the quantization of the full system, as already mentioned. There will thus arise many new issues, such as for example the link between a formally integrable system in infinitely many variables, and the chaoticity of a finite-dimensional system obtained from it by projection to finitely many variables (some comments on this issue can be found in [7]).

6.5 Supersymmetry

The quantum billiard analysis can be extended to maximal *super*gravity, with $d=10$, by supplementing the bosonic degrees of freedom with a vector-spinor, the gravitino. In the E_{10} approach, the latter corresponds to a spinorial representation of the "R-symmetry" group $K(E_{10})$ in terms of which the Rarita–Schwinger equation of $D=11$ supergravity (with the usual truncations) can be rewritten as a $K(E_{10})$ covariant "Dirac equation" [44–46]. When restricting to the diagonal metric degrees of freedom, the gravitino ψ_μ of $D=11$ supergravity performs a separate classical[6] fermionic billiard motion [49]. This is most easily expressed in a supersymmetry gauge $\psi_t = \Gamma_t \Gamma^a \psi_a$ [46] and in the variables [49] (with $\Gamma_* = \Gamma^1 \cdots \Gamma^{10}$):

$$\varphi^a = g^{1/4} \Gamma_* \Gamma^a \psi^a \quad \text{(no sum on } a = 1, \ldots, 10) \tag{6.28}$$

(recall that $g \equiv \exp(-2\sum \beta^a)$). Using (6.28) in conjunction with equation (6.3) of [48], the Dirac brackets between two gravitino variables become

$$\left\{\varphi^a_\alpha, \varphi^b_\beta\right\}_{\text{D.B.}} = -2i\, G^{ab} \delta_{\alpha\beta} \tag{6.29}$$

where we have written out the 32 real spinor components using the indices α, β. We stress that it is precisely the inverse DeWitt metric G^{ab}, see (6.3), which appears in this equation!

The fermionic and bosonic variables are linked by the supersymmetry constraint

$$\mathcal{S}_\alpha \equiv \sum_{a,b=1}^{10} \dot{\beta}^a G_{ab} \varphi^b_\alpha = \sum_{a=1}^{10} \pi_a \varphi^a_\alpha = 0\,. \tag{6.30}$$

The supersymmetry constraint implies the Hamiltonian constraint $\mathcal{H}_0 = 0$ by closure of the algebra

$$\frac{1}{4}\left\{\mathcal{S}_\alpha, \mathcal{S}_\beta\right\}_{\text{D.B.}} = -i\,\delta_{\alpha\beta} \mathcal{H}_0\,. \tag{6.31}$$

6 In the sense that the gravitino is treated as a classical variable, not as an operator.

In order to quantize this system we rewrite the 320 real gravitino components φ^a_α in terms of 160 complex ones according to

$$\widetilde{\varphi}^a_A := \varphi^a_A + i\varphi^a_{A+16} \tag{6.32}$$

for $A, B, \ldots = 1, \ldots, 16$, and replace the Dirac brackets (6.29) by canonical anticommutators

$$\left\{\widetilde{\varphi}^a_A, (\widetilde{\varphi}^\dagger)^b_B\right\} = 2G^{ab}\delta_{AB}, \quad \left\{\widetilde{\varphi}^a_A, \widetilde{\varphi}^b_B\right\} = \left\{(\widetilde{\varphi}^\dagger)^a_A, (\widetilde{\varphi}^\dagger)^b_B\right\} = 0 \tag{6.33}$$

to obtain a fermionic Fock space of dimension 2^{160} by application of the fermionic creation operators $(\widetilde{\varphi}^a_A)^\dagger$ to the vacuum $|\Omega\rangle$ defined by

$$\widetilde{\varphi}^a_A |\Omega\rangle = 0. \tag{6.34}$$

For the supersymmetry constraint this amounts to the redefinition $\widetilde{\mathcal{S}}_A = \mathcal{S}_A + i\mathcal{S}_{A+16}$. Because (6.34) implies $\widetilde{\mathcal{S}}_A|\Omega\rangle = 0$, the quantum supersymmetry constraint is then solved by

$$|\Psi\rangle = \prod_{A=1}^{16} \widetilde{\mathcal{S}}^\dagger_A \Big(\Phi(\rho, z)|\Omega\rangle\Big), \tag{6.35}$$

if and only if the function $\Phi(\rho, z)$ is a solution of the WDW equation $\mathcal{H}_0\Phi = 0$. While this solution is close to the "bottom of the Dirac sea," there is an analogous one "close to the top" with $\widetilde{\mathcal{S}}^\dagger_A$ replaced by $\widetilde{\mathcal{S}}_A$ and $|\Omega\rangle$ by the completely filled state.

In existing studies of the fermionic sector so far the fermions have been treated "quasi-classically," that is, as Grassmann-valued c-numbers. However, the correspondence between the full supergravity equations of motion and the fermionic extension of the $E_{10}/K(E_{10})$ model, as far as it has been established, is lacking inasmuch as the relevant spinorial representations of the "R-symmetry" group $K(E_{10})$ identified to date are all *unfaithful*, and hence cannot capture the full fermionic dynamics of M-theory (see [48] for a detailed discussion of this problem). Again, quantization may be essential here; in fact, a satisfactory solution and incorporation of all the fermionic degrees of freedom into the E_{10} model may require some kind of "bosonization" of the fermionic degrees of freedom. This would also be in accord with the fact that fermions are intrinsically quantum objects.

6.6 Outlook

The cosmological billiards description takes into account the dependence on spatial inhomogeneities and matter degrees of freedom only in a very rudimentary way via the infinite potential walls. It would thus be desirable to develop an approximation scheme for the quantum state in line with the "small tension" expansion proposed in [2], and thereby hopefully resolve the difficulties encountered in extending the "dictionary" of [2] to higher-order spatial gradients and heights of roots in a quantum mechanical context. In the strict BKL approximation, the full wavefunction is expected to factorize as (see also [50])

$$|\Psi_{\text{full}}\rangle \sim \prod_{\mathbf{x}} |\Psi_{\mathbf{x}}\rangle\,, \tag{6.36}$$

near the singularity into a formal ultralocal product over wavefunctions of type (6.35), one for each spatial point, with independent bosonic wavefunctions $\Phi_{\mathbf{x}}\big(\rho(\mathbf{x}), z(\mathbf{x})\big)$ and *space-dependent* metric variables $\beta^a(\mathbf{x}) \equiv (\rho(\mathbf{x}), z(\mathbf{x}))$. At first sight, it may seem paradoxical that (6.36) should become a better and better approximation near the singularity, as all the dynamics gets concentrated in a continuous superposition of Cartan subalgebras, whereas one would expect the full tower of E_{10} Lie algebra states, rather than just the Cartan subalgebra degrees of freedom, to become excited – in analogy with string cosmology, where one would expect the full tower of string states to become relevant near the singularity. However, the apparent paradox may well turn out to be the crux of the matter: the task is to replace the formal expression (6.36) involving a formal continuous product over all spatial points by a wavefunction which depends solely on the (infinite) tower of E_{10} degrees of freedom, and where all spatial dependence is discarded. It is this step which would effectively implement the de-emergence of space and time near the cosmological singularity, and their replacement by purely algebraic concepts [46, 51].

For this purpose we will have to generalize the minisuperspace Hamiltonian (6.9) to the full E_{10} Lie algebra. In fact, as a consequence of the uniqueness of the quadratic Casimir operator on E_{10}, there is a *unique* E_{10} extension of the billiard Hamiltonian (6.9) given by

$$\mathcal{H}_0 \to \mathcal{H} = \mathcal{H}_0 + \sum_{\alpha \in \Delta_+(E_{10})} \sum_{s=1}^{\text{mult}(\alpha)} e^{-2\alpha(\beta)} \Pi^2_{\alpha,s}\,, \tag{6.37}$$

where the first sum runs over the positive roots α of E_{10} and the second one over a basis of the possibly degenerate root space of α. Due to our lack of a manageable realization of the E_{10} algebra, this is a highly formal expression, but we can nevertheless note two important features: like the minisuperspace Hamiltonian (6.9), this operator is free of ordering ambiguities by the uniqueness of the E_{10} Casimir operator, and it has the form of a *free* Klein–Gordon operator, albeit in infinitely many dimensions. The absence of potential terms is due to the fact that in the approach of [2] (as far as it has been worked out, at least) the spatial gradients, which give rise to the "potential" $\propto R^{(3)}$ in the standard WDW equation, are here replaced by *time derivatives of dual fields*. Accordingly, the contributions to the potential are replaced by momentum-like operators $\propto \Pi^2$. It is for this reason that the restriction to positive-norm wavefunctions may still be consistent for the full E_{10} theory – unlike for the standard WDW equation (but we note that so far no one has succeeded in generalizing the minisuperspace scalar product (6.11) to the full theory).

Nevertheless, the Hamiltonian (6.37) is not the complete story because, as in standard canonical gravity, the extended system requires additional constraints. For the E_{10} model their complete form is not known, but a first step towards their incorporation was taken in [52], where a correspondence was established at low levels between the classical canonical constraints of $D = 11$ supergravity (in particular, the diffeomorphism and Gauss constraints) on the one hand, and a set of constraints that can be consistently imposed on the $E_{10}/K(E_{10})$ coset space dynamics on the other (see [53] for more recent results in this direction). The fact that the latter can be cast in a "Sugawara-like" form quadratic in the E_{10} Noether charges [52] would make them particularly amenable for the implementation on a quantum wavefunction. In addition, one would expect that $PSL_2(\mathtt{O})$ must be replaced by a much larger "modular group" whose action extends beyond the Cartan subalgebra degrees of freedom all the way into E_{10}, perhaps along the lines suggested in [54].

As noted above, the inequality (6.20) implies that $\Phi(\rho, z) \to 0$ for $\rho \to \infty$, and hence the wavefunction Ψ vanishes at the singularity, in such a way that the norm is preserved. Its oscillatory nature entails that it cannot be analytically extended beyond the singularity, a result whose implications for the question of singularity resolution in quantum cosmology remain to be explored. At least formally, these conclusions remain valid in the full theory: the extra contribution in (6.37) extending $\mathcal{H}_0$ to the full Hamiltonian $\mathcal{H}$ is positive, as follows from the manifest positivity of the E_{10} Casimir operator on the complement of the Cartan subalgebra of E_{10}. Hence the inequality (6.20) is further strengthened and thus constitutes a lower bound also for the full Hamiltonian.

To put our results in perspective, we recall that the mechanism usually invoked to resolve singularities in canonical approaches to quantum geometrodynamics would be to replace the classical "trajectory" in the moduli space of 3-geometries (that is, WDW superspace) by a quantum mechanical wavefunctional which "smears" over the singular 3-geometries. By contrast, the present work suggests a very different picture, namely the "resolution" of the singularity via the *effective disappearance (de-emergence) of space-time* near the singularity (see also [51]). The singularity would thus become effectively "unreachable." This behavior is very different from other possible mechanisms, such as the Hartle–Hawking no-boundary proposal

[55], or cosmic bounce scenarios of the type considered recently in the context of minisuperspace loop quantum cosmology [56–58], both of which require continuing the cosmic wavepacket into and beyond the singularity at $\rho = \infty$. In contrast to these models, the exponentially growing complexity of the E_{10} Lie algebra suggests that it may turn out to be impossible to "resolve" the quantum equations as one gets closer and closer to the singularity. In other words, there may appear an element of *non-computability* (in a mathematically precise sense) that may forever screen the Big Bang from complete resolution.

A key question for singularity resolution concerns the role of observables, and their behavior near the singularity. While no observables (in the sense of Dirac) are known for canonical gravity, we only remark here that for the $E_{10}/K(E_{10})$ coset model the conserved E_{10} Noether charges do constitute an infinite set of observables, as these charges can be shown to commute with the full E_{10} Hamiltonian (6.37). The expectation values of these charges are the only quantities that remain well-defined and can be sensibly computed in the deep quantum regime, where the $E_{10}/K(E_{10})$ coset model is expected to replace space-time-based quantum field theory. In the final analysis these charges would thus replace geometric quantities (such as the curvature scalar) which blow up at the singularity, but which are not canonical observables.

Acknowledgments

H. N. is grateful to the Institute for Advanced Studies in Stellenbosch, South Africa, and especially D. Loureiro, J. Murugan, and A. Weltman for hospitality and for organizing a wonderful conference in honor of George Ellis' 70th birthday. We thank J. Barbour, M. Berry, M. Koehn, R. Penrose, and H. Then for discussions and correspondence. A. K. is a Research Associate of the Fonds de la Recherche Scientifique–FNRS, Belgium.

References

[1] A. Kleinschmidt, M. Koehn, and H. Nicolai, *Phys. Rev.* **D80** (2009) 061701(R) [arXiv:0907.3048 [gr-qc]].

[2] T. Damour, M. Henneaux, and H. Nicolai, *Phys. Rev. Lett.* **89** (2002) 221601 [arXiv:hep-th/0207267].

[3] E. Cremmer and B. Julia, *Nucl. Phys. B* **159** (1979) 141.

[4] V. A. Belinskii, I. M. Khalatnikov, and E. M. Lifshitz, *Adv. Phys.* **19** (1970) 525; *Adv. Phys.* **31** (1982) 639.

[5] C. W. Misner, in *Deterministic Chaos in General Relativity*, edited by D. Hobill *et al.* (Plenum Press, New York, 1994).

[6] M. P. Ryan and L. C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton University Press, New Jersey, 1975).

[7] T. Damour, M. Henneaux, and H. Nicolai, *Class. Quant. Grav.* **20** (2003) R145 [arXiv:hep-th/0212256].
[8] M. Henneaux, D. Persson, and P. Spindel, *Living Rev. Rel.* **11** (2008) 1 [arXiv:0710.1818 [hep-th]].
[9] T. Damour and M. Henneaux, *Phys. Rev. Lett.* **86** (2001) 4749 [arXiv:hep-th/0012172].
[10] T. Damour, M. Henneaux, B. Julia, and H. Nicolai, *Phys. Lett.* **B509** (2001) 323 [arXiv: hep-th/0103094].
[11] B. Julia, in *Lectures in Applied Mathematics*, Vol. 21, AMS-SIAM (1985) 335, LPTENS-82-22
[12] B. S. DeWitt, *Phys. Rev.* **160** (1967) 1113.
[13] C. W. Misner, *Phys. Rev. Lett.* **22** (1969) 1071.
[14] C. W. Misner, in *Magic Without Magic*, edited by J. R. Klauder (W.H. Freeman & Co, San Francisco, 1972) 441–473.
[15] A. Macias, O. Obregon, and M. P. Ryan, *Class. Quant. Grav.* **4** (1987) 1477.
[16] R. Graham and P. Szepfalusy, *Phys. Rev.* **D42** (1990) 2483.
[17] V. D. Ivanchuk and V. N. Melnikov, *Class. Quant. Grav.* **12** (1995) 809, arXiv:gr-qc/9407028.
[18] P. V. Moniz, *Int. J. Mod. Phys.* A **11** (1996) 4321 [arXiv:gr-qc/9604025].
[19] P. D. D'Eath, *Supersymmetric Quantum Cosmology* (Cambridge University Press, Cambridge, 1996).
[20] A. Csordas and R. Graham, *Phys. Rev. Lett.* **74** (1995) 4129 [arXiv:gr-qc/9502004].
[21] C. Kiefer, *Quantum Gravity* (Clarendon Press, Oxford, 2004).
[22] L. A. Forte, *Class. Quant. Grav.* **26** (2009) 045001 [arXiv:0812.4382 [gr-qc]].
[23] B. Pioline and A. Waldron, *Phys. Rev. Lett.* **90** (2003) 031302 [arXiv:hep-th/0209044].
[24] D. Giulini and C. Kiefer, *Phys. Lett.* **A193** (1994) 21.
[25] H. Iwaniec, *Spectral Methods of Automorphic Forms*, American Mathematical Society Graduate Studies in Mathematics, Vol. 53 (2002).
[26] C. Isham, in *Recent Aspects of Quantum Fields*, edited by H. Mitter and H. Gausterer, Springer Lecture Notes in Physics 396 (1991).
[27] J. Barbour, *Phys. Rev. D* **47** (1993) 5422.
[28] J. Barbour, *The End of Time*, Weidenfeld & Nicolson, Orion Publishing Group, London (1999).
[29] A. J. Feingold, A. Kleinschmidt, and H. Nicolai, *J. Algebra* **322** (2009) 1295 [arXiv:0805.3018 [math.RT]].
[30] H. M. S. Coxeter, *Duke Math. J.* **13** (1946) 561.
[31] J.H. Conway and D.A. Smith, *On Quaternions and Octonions* (A.K. Peters, Wellesley, MA, 2003).
[32] M. Eie and A. Krieg, *Math. Z.* **210** (1992) 113.
[33] A. Kleinschmidt, H. Niedai and J. Palmkvist, arXiv 1010.2212 [math.NT].
[34] G. Steil, DESY Preprint (1994) DESY 94-028.
[35] D. A. Hejhal, in *Emerging Applications of Number Theory*, edited by D. A. Hejhal *et al.*, IMA Series No. 109 Springer (1999) 291–315.
[36] For a review of quantum and arithmetical chaos, see: E. Bogomolny, nlin/0312061.
[37] H. Then, *Math. Comp.* **74** (2004) 363 [math-ph/0305047].
[38] R. Aurich, F. Steiner, and H. Then, arXiv:gr-qc/0404020.
[39] M. V. Berry, *J. Phys. A* **10** (1977) 2083.
[40] M. V. Berry, *Ann. Phys.* **131** (1981) 163.
[41] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer, New York, 1990).
[42] M. Sieber and F. Steiner, *Phys. D* **44** (1990) 248.

[43] E. J. Heller, *Phys. Rev. Lett.* **53** (1984) 1515.
[44] M. Berry and R.J. Mondragon, *Proc. R. Soc. Lond.* **A412** (1987) 53.
[45] M. J. Davis and E. J. Heller, *J. Chem. Phys.* **75** (1981) 3916.
[46] T. Damour, A. Kleinschmidt, and H. Nicolai, *Phys. Lett. B* **634** (2006) 319 [arXiv: hep-th/0512163].
[47] S. de Buyl, M. Henneaux, and L. Paulot, JHEP **0602** (2006) 056 [arXiv: hep-th/0512292].
[48] T. Damour, A. Kleinschmidt, and H. Nicolai, JHEP **0608** (2006) 046 [arXiv:hep-th/0606105].
[49] T. Damour and C. Hillmann, JHEP **0908** (2009) 100 arXiv:0906.3116 [hep-th].
[50] A. A. Kirillov, *JETP Lett.* **62** (1995) 89 [*Pisma Zh. Eksp. Teor. Fiz.* **62** (1995) 81].
[51] T. Damour and H. Nicolai, *Int. J. Mod. Phys. D* **17** (2008) 525 [arXiv:0705.2643 [hep-th]].
[52] T. Damour, A. Kleinschmidt, and H. Nicolai, *Class. Quant. Grav.* **24** (2007) 6097 [arXiv:0709.2691 [hep-th]].
[53] T. Damour, A. Kleinschmidt, and H. Nicolai, arXiv:0912.3491 [hep-th].
[54] O. Ganor, arXiv:hep-th/9903110.
[55] J. B. Hartle and S. W. Hawking, *Phys. Rev. D* **28** (1983) 2960.
[56] M. Bojowald, *Phys. Rev. Lett.* **86** (2001) 5227 [arXiv:gr-qc/0102069].
[57] A. Ashtekar, A. Corichi, and P. Singh, *Phys. Rev. D* **77** (2008) 024046 [arXiv:0710.3565 [gr-qc]].
[58] M. Bojowald, *Phys. Rev. Lett.* **100** (2008) 221301 [arXiv:0805.1192 [gr-qc]].

7

Progress in RNS string theory and pure spinors

DIMITRY POLYAKOV

This chapter is a review of the program of understanding the gauge theory/gravity correspondence through a study of the RNS superstring. In particular, we show how to define string field theory actions in curved backgrounds by constructing a sequence of new nilpotent BRST operators in RNS string theory. Our construction is based on the presence of local gauge symmetries in RNS superstring theory leading to an infinite chain of new BRST generators that can be classified in terms of ghost cohomologies.

7.1 Introduction

Gauge–string duality is arguably one of the most profound problems in modern physics [1–11]. This duality implies that the gauge-invariant observables (operators) in QCD are in one-to-one correspondence with the physical states (vertex operators) in string theory. The reason why extended objects (such as strings) appear in QCD is quite natural. If we recall standard electrodynamics, there are two ways of describing it: either in terms of local electric fields (Coulomb's approach), or in terms of the geometry of electric field lines (Faraday's approach). In the case of electromagnetic theory, Coulomb's approach turns out to be far more efficient. In the case of QCD, however, things are quite different. While the electric field lines created by the charged particles are spread over the entire space, the gluon field lines are confined to thin flux tubes. These flux tubes, connecting quarks, can be naturally interpreted as one-dimensional extended objects, known as QCD strings. Thus the interaction of quarks in QCD could naturally be described by the dynamics of flux lines (QCD strings) connecting the quarks. Such a description would

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

provide a natural solution to the problem of confinement, which is a long-standing problem in QCD of crucial importance. Indeed, the confinement of quarks (the fact that they exist only in bound states, cannot be separated and we can't observe them individually) implies that the interaction energy between quarks grows with the distance. Such a model can be realised, for example, in the scenario of quarks connected by 'springs', represented by the gluon flux tubes, described above. More precisely, the well-known confinement criterion, based on the condition that the interaction energy between quarks grows linearly with the distance, can be reformulated naturally in the language of the Wilson loops. Namely, consider a process in which one attempts to take apart a pair of quarks (say, a quark–antiquark pair constituting a meson), starting from a certain point A in the space. The quark and the antiquark are then brought together at a point B. The whole process is then described by the closed contour C, passing through A and B. Then the confinement criterion (equivalent to the condition of the linear energy growth with the distance) is that the expectation value of the Wilson loop decays exponentially with the area:

$$W(C) = \left\langle Tr\, P e^{\oint_C A_n(X(s))dX^m(s)} \right\rangle \sim e^{-A(C)}, \tag{7.1}$$

where $0 \leq s \leq 2\pi$ is the parameter of the loop. The gauge–string correspondence then implies that $W(C)$ is identified with the partition function $Z(C)$ of an open string with the ends attached to the same contour C. The expansion of $W(C)$ in terms of local gauge-invariant operators corresponds to the expansion of $Z(C)$ in terms of physical string-theoretic vertex operators. Since, by definition, the action of a string is given by the area of its worldsurface, then $Z(C)$, at least in a WKB approximation, is given by [2, 11]:

$$Z(C) = \int e^{-S_{\text{string}}(C)} \sim e^{-A_{\min}(C)}, \tag{7.2}$$

where $A_{\min}(C)$ is the area of the minimal surface spanned by C. Therefore, if one is able to find a string theory whose physical degrees of freedom match those of QCD, one could expect that correlation functions of massless vertex operators in open string theory will reproduce QCD dynamics. Such a string-theoretic framework would be a particularly efficient and natural way to address the problem of confinement, as well as other non-perturbative QCD dynamics. In practice, however, things are far more complicated. First of all, if an open string is to describe the gluon dynamics, its spectrum has to contain eight massless vector bosons. It is well known that a perturbative open string spectrum has only one massless physical excitation, a photon, which has no colour. Although this complication can be corrected (albeit somewhat artificially) by introducing the appropriate Chan–Paton factors [12] (such as Gell-Mann matrices), this is only the beginning of our troubles.

The main and fundamental problem in identifying a QCD string [2–5], is that a spectrum of a normal open string contains an infinite tower of massive states, in addition to the photon. These states appear as intermediate poles in any Veneziano amplitude (including the scattering of massless gauge bosons), so there is no way to separate the gluon dynamics from massive string modes. For this reason, the Veneziano amplitude for the massless gauge bosons in superstring theory has little to do with the scattering amplitudes of gluons in QCD, since the latter do not, of course, have anything like an infinite set of intermediate massive states. This complication has an underlying geometrical reason. That is, the standard open string theory lacks an important symmetry that, however, is present on the gauge theory side. Namely, while the string theory is only invariant under reparametrisations with positive Jacobians (that do not change the worldsheet orientation), the Wilson loop is also invariant under the orientation change. Technically, the loop equation satisfied by $\langle W(C)\rangle$, is the consequence of the zigzag symmetry [5]. It is easy to check that Ward identities for perturbative string theory in flat space-time background do not reproduce the loop equation, indicating the absence of the zigzag symmetry. In other words, a stringy description of QCD requires the presence of a closed subalgebra of massless vertex operators of gluons and, as a consequence, the field-theoretic behaviour of their scattering amplitudes (including the absence of the intermediate poles in correlation functions and OPE). Finding the string description of QCD thus appears to be a challenging and difficult problem.

One well-known example of gauge–string duality, AdS/CFT, is too dependent on SUSY arguments [3–8]. Certainly the $N=2$ or $N=4$ supersymmetries it relies on probably do not exist in reality. One possibility (and the main theme of this chapter) is that in order to understand the *non-SUSY* version of gauge–string duality, we need to elevate the language in which we describe both QCD and string theory. On the QCD side this means that, instead of describing the theory in terms of local observables (gauge fields) living in usual geometrical space, we shall describe it in terms of extended objects – the gluon flux lines connecting quarks, that live in infinite-dimensional 'loop space' [1]. The loop equations in the large-N limit are:

$$\hat{L}(s)W(C)=0, \tag{7.3}$$

where $W(C)$ is the expectation value of the Wilson loop and the operator

$$\hat{L}(s)=\lim_{\varepsilon\to 0}\int_{-\varepsilon}^{\varepsilon} d\alpha \frac{\delta^2}{\delta X_m(s-\alpha)\delta X^m(s+\alpha)}$$

is the functional Laplacian defined in the loop space. This reproduces the YM equations on the classical level [1]:

$$\nabla_m F^{mn}=0 \tag{7.4}$$

and the complete set of the QCD Feynman diagrams in the quantum theory. The loop equation fully encodes the non-perturbative dynamics of QCD and, for contours without self-intersections, can be formally obtained from the action in the loop space:

$$S \sim \int_C W(C)\hat{L}W(C)$$

which structurally is similar to the free part of the SFT action:

$$S \sim \int \Psi Q_{\text{brst}} \Psi$$

(Ψ here is a string field comprising all the off-shell vertex operators). Since the Wilson loop can be expanded in terms of the local gauge-invariant operators Makeenko and Migdal (1979):

$$W(C) = 1 + \frac{1}{2}a^2 Tr\left(F^2\right) + O\left(F^4\right) + \ldots$$

and, according to the gauge–string duality conjecture all such operators correspond to vertex operators in the dual closed string theory on AdS_5, e.g.

$$Tr\left(F^2\right) \leftrightarrow V_{\text{dilaton}}.$$

The loop equations imply that:

- On the gauge theory side, all the local gauge-invariant observables are annihilated by $\hat{L}$.
- The gauge–string duality implies $\hat{L} \leftrightarrow Q_{\text{brst}}$.
- This means the loop equation $\hat{L}W(C) = 0$ simply corresponds to the equations of motion of string field theory: $Q_{\text{brst}}\Psi = 0$ [13–12] in the dual theory.
- Therefore, in order to understand the non-SUSY gauge–string duality it is crucial to understand the dynamics of SFT built around *curved* backgrounds.
- However, in standard approaches SFT has so far been developed around *flat* background only; extending the existing construction to curved geometries appears to be a hard problem [19].

In this chapter we outline the program in this direction by constructing a sequence of new nilpotent BRST operators in RNS string theory [21, 22] defining the SFT actions in curved backgrounds. Our construction [23–26] is based on the presence of local gauge symmetries in RNS superstring theory (that we point out to exist), leading to an infinite chain of new BRST generators that can be classified in terms of ghost cohomologies.

These gauge symmetries are closely related to global non-linear space-time α-symmetries in RNS superstring theory that mix matter and ghost degrees of freedom, form ground rings and originate from hidden space-time dimensions. The new nilpotent BRST charges, whose construction will be demonstrated in this chapter, correspond to a sequence of RNS superstring theories in curved backgrounds (including AdS-type) and can be used to develop SFTs around non-trivial backgrounds. In terms of RNS–pure spinor (PS) correspondence, we show that the appearance of new BRST charges corresponds to introducing interactions for the pure spinor variable λ^α in the PS BRST operator $\oint \frac{dz}{2i\pi}\lambda^\alpha d_\alpha$ (equivalent to OPE singularities between λ's that preserve the nilpotence of Q_{BRST}). The orders of ghost cohomologies of BRST charges in RNS formalism correspond to the leading order of OPE singularity of two pure spinors in pure spinor formalism.

In string theory the global space-time symmetries are typically generated by primary fields of conformal dimension 1 (commuting with BRST charge), while local gauge symmetries are given by BRST exact operators (of various conformal dimensions and not necessarily primary), given by commutators of BRST operators with appropriate ghost fields. Examples of generators of local gauge symmetries on the worldsheet are the stress–energy tensor T and the supercurrent G:

$$T = \{Q_0, b\} \tag{7.5}$$

and

$$G = [Q_0, \beta], \tag{7.6}$$

where

$$Q_0 = \oint \frac{dz}{2i\pi}\left(cT + \partial ccb - \frac{1}{2}\gamma\psi_m\partial X^m - \frac{1}{4}b\gamma^2\right)$$

is the standard BRST charge. At the same time, integrals of conformal dimension 1 BRST non-trivial primary fields:

$$L^m = \oint \frac{dz}{2i\pi}\partial X^m \tag{7.7}$$

and

$$L^{mn} = \oint \frac{dz}{2i\pi}\psi^m\psi^n \tag{7.8}$$

define Lorentz translations and rotations in d-dimensional space-time. The rotation generator $L^{mn} = \oint \frac{dz}{2i\pi}\psi^m\psi^n$ is incomplete, as it acts only on the ψ's but not on the X's. This is due to the BRST non-invariance of L^{mn} (it commutes only with the stress tensor part of Q_0, but not with the supercurrent terms). To construct a complete version of the generator of Lorentz rotations, which acts both on X's and ψ's, one needs to improve L^{mn} with bc ghost-dependent terms. The construction of the complete version of L^{mn} can be done using the K-operator prescription [23]:

- Let $L = \oint V$ be some non-invariant operator satisfying

$$[Q_0, V] = \partial U + W$$

so that

$$[Q_0, L] = W.$$

- Consider the operator

$$K = -4ce^{2\chi - 2\phi} \equiv \xi \Gamma^{-1}$$

satisfying

$$\{Q_0, K\} = 1.$$

Then, if the operator product of K and W is non-singular, satisfying

$$K(z)L(w) \sim O((z-w)^N)$$

$(N > 0)$ the invariant generator $\widetilde{L}$ can be obtained from non-invariant $L = \oint V$ by the K-transformation:

$$\begin{aligned}\widetilde{L}(w) = L + \frac{(-1)^{N+1}}{N!} \oint \frac{dz}{2i\pi}(z-w)^N : K\partial_z^N W(z) \\ + \frac{1}{N!}\oint \frac{dz}{2i\pi}\partial_z^{N+1}\left[(z-w)^N K\right] K\{Q_0, U\}(z).\end{aligned}$$

Note the dependence on an arbitrary worldsheet point w in all but the $N = 0$ case! For the rotation generator L^{mn}, $N = 0$ and its complete version is given by

$$\widetilde{L^{mn}} = L^{mn} - 2\oint \frac{dz}{2i\pi} c\xi e^{-\phi}\partial X^{[m}\psi^{n]} - \frac{1}{2}\partial c c e^{3\chi - 3\phi}\partial X^{[m}\psi^{n]}$$

with the bosonic and fermionic ghosts β, γ, b, c bosonised as

$$\begin{aligned}&\gamma(z) = e^{\phi - \chi}(z);\\ &\beta(z) = e^{\chi - \phi}\partial\chi(z) \equiv \partial\xi e^{-\phi}(z);\\ &b(z) = e^{-\sigma}(z);\, c(z) = e^{\sigma}(z).\end{aligned}$$

This defines the BRST-invariant rotation generator, acting both on ψ's and (up to a picture-changing) on X's. The next, far less trivial example of global space-time supersymmetry in superstring theory is given by the hierarchy of α-symmetries. These global space-time symmetries are realised non-linearly, mixing the matter and the ghost sectors of RNS superstring theory, and can be classified in terms of ghost cohomologies. Namely, it can be checked that the full matter+ghost RNS action:

$$\begin{aligned}
S_{\mathrm{RNS}} &= S_{\mathrm{matter}} + S_{bc} + S_{\beta\gamma} \\
S_{\mathrm{matter}} &= \frac{1}{2\pi}\int d^2z(\partial X_m \bar{\partial} X^m + \psi_m \bar{\partial}\psi^m + \bar{\psi}_m \partial \bar{\psi}^m) \\
S_{bc} &= \frac{1}{2\pi}\int d^2z(b\bar{\partial}c + \bar{b}\partial\bar{c}) \\
S_{\beta\gamma} &= \frac{1}{2\pi}\int d^2z(\beta\bar{\partial}\gamma + \bar{\beta}\partial\bar{\gamma})
\end{aligned}$$

is invariant under the following transformations (with α being a global parameter):

$$\begin{aligned}
\delta X^m &= \alpha\{2e^{\phi}\partial\psi^m + \partial(e^{\phi}\psi^m)\} \\
\delta\psi^m &= -\alpha\{e^{\phi}\partial^2 X^m + 2\partial(e^{\phi}\partial X^m)\} \\
\delta\gamma &= \alpha e^{2\phi-\chi}\{\psi_m\partial^2 X^m - 2\partial\psi_m\partial X^m\} \\
\delta\beta &= \delta b = \delta c = 0.
\end{aligned}$$

This of course means that

$$\begin{aligned}
\delta S_{\mathrm{matter}} &= -\delta S_{\beta\gamma} = \frac{1}{2\pi}\int d^2z(\bar{\partial}e^{\phi})(\psi_m\partial^2 X^m - 2\partial\psi_m\partial X^m) \\
\delta S_{bc} &= \delta S_{\mathrm{RNS}} = 0.
\end{aligned}$$

The generator of these transformations is given by

$$L^{\alpha} = \oint \frac{dz}{2i\pi} e^{\phi} F(X,\psi) \equiv \oint \frac{dz}{2i\pi} e^{\phi}(\psi_m\partial^2 X^m - 2\partial\psi_m\partial X^m),$$

where it is convenient to introduce the notation for the dimension $\frac{5}{2}$ primary field:

$$F(X,\psi) = \psi_m\partial^2 X^m - 2\partial\psi_m\partial X^m$$

along with the matter worldsheet supercurrent $G = -\frac{1}{2}\psi_m\partial X^m$ and the dimension 2 primary $L(X,\psi) = 2\partial\psi_m\psi^m - \partial X_m\partial X^m$, which is the w.s. superpartner of F, i.e. $G(z)L(w) \sim \frac{F(w)}{z-w}$. The L^{α}-generator is an element of the ghost cohomology H_1. As in the case of the rotation generator, the integrand of the L^α-generator is a primary field of dimension 1, however, it is not BRST-invariant since it doesn't commute with the supercurrent terms of the BRST charge. So similarly to L^{mn}-rotations one has to apply the K-transformation to make it BRST-invariant. For the α-symmetry generator, however, the K-operator procedure gives $N=2$! Positive ghost cohomologies H_n with $n>0$ consist of picture-inequivalent physical operators, existing at pictures n and above, annihilated by inverse picture-changing transformation at minimal positive picture n. Negative ghost cohomologies H_{-n} consist of picture-inequivalent physical operators, existing at pictures $-n$ and below, annihilated by direct picture-changing at minimal negative picture $-n$.

An isomorphism holds between positive and negative cohomologies: $H_n \sim H_{-n-2}$. H_0 by definition consists of picture-equivalent operators existing at all pictures (including picture 0), while H_{-1} and H_{-2} are empty. The full BRST-invariant extension of L^α generating the complete set of α-symmetries for the matter and the ghost sectors is given by [23–26]:

$$\widetilde{L}^\alpha(w)=\oint\frac{dz}{2i\pi}(z-w)^2\Big\{e^\phi F P^{(2)}_{2\phi-2\chi-\sigma}(z)+8c\xi\left(FG-\frac{1}{2}LP^{(2)}_{\phi-\chi}-\frac{1}{4}\partial LP^{(1)}_{\phi-\chi}\right)$$
$$-24\partial cce^{2\chi-\phi}F\Big\}\equiv\oint\frac{dz}{2i\pi}(z-w)^2V_3,$$

where the conformal weight n polynomials $P^{(n)}_{f(\phi_1(z),\ldots,\phi_N(z))}$ are the generalised Hermite polynomials defined as

$$P^{(n)}_{f(\phi_1(z),\ldots,\phi_N(z))}=e^{-f(\phi_1(z),\ldots,\phi_N(z))}\frac{\partial^n}{\partial z^n}e^{f(\phi_1(z),\ldots,\phi_N(z))}$$

for an arbitrary function f of N fields $\phi_1(z),\ldots,\phi_N(z)$ (e.g. $P^{(1)}_{2\phi-2\chi-\sigma}=2\partial\phi-2\partial\chi-\partial\sigma$). The operator $\widetilde{L}^\alpha$ is BRST-invariant and non-trivial, generating the full set of global non-linear space-time symmetries, originating from hidden dimensions. Note that the dimension 3 integrand of $\widetilde{L^\alpha}$ satisfies

$$[Q_0,V_3]=\partial^3W_0$$

where W_0 is a dimension 0 operator (whose precise form is skipped for brevity). While it depends on an arbitrary worldsheet coordinate w, this dependence doesn't affect any correlation functions, as the w derivatives of $\widetilde{L}^\alpha(w)$ are BRST exact, forming the ground ring.

The non-vanishing operators are the first and second derivatives of $\widetilde{L}^{\alpha}(w)$ in w, given by

$$L_1^{\alpha}(w) = \partial_w L^{\alpha}(w)$$
$$= -2\oint \frac{dz}{2i\pi}(z-w)\Big\{e^{\phi} F P^{(2)}_{2\phi-2\chi-\sigma}(z) + 8c\xi\left(FG - \frac{1}{2}LP^{(2)}_{\phi-\chi} - \frac{1}{4}\partial L P^{(1)}_{\phi-\chi}\right)$$
$$-24\partial c c e^{2\chi-\phi}F\Big\}$$

and

$$L_2^{\alpha}(w) = \partial_w L_1^{\alpha}(w).$$

It is straightforward to show that these generators induce local gauge symmetries on the worldsheet and are BRST exact, that is:

$$\begin{aligned} L_1^{\alpha}(w) &= \left\{Q_0, \oint \frac{dz}{2i\pi}\widetilde{b}(z,w)\right\} \\ L_2^{\alpha}(w) &= \left\{Q_0, \partial_w \oint \frac{dz}{2i\pi}\widetilde{b}(z,w)\right\} \end{aligned} \tag{7.9}$$

with the role of the generalised b-ghost (corresponding to gauge transformations induced by $L_j^{\alpha} \equiv \partial_w^j L^{\alpha}(w)$; $j = 1, 2$) played by

$$\oint \frac{dz}{2i\pi}\widetilde{b}(z,w) = \oint \frac{dz}{2i\pi}(z-w)^2\Big\{-2be^{\phi}FP^{(1)}_{2\phi-2\chi-\sigma}(z)$$
$$+8\xi\left(FG - \frac{1}{2}LP^{(2)}_{\phi-\chi} - \frac{1}{4}\partial L P^{(1)}_{\phi-\chi}\right) + 24\partial c e^{2\chi-\phi}F\Big\}.$$

Notice that the integrands of L_1^{α} and $\widetilde{b}$ are conformal dimension 2 generators. Now that we have the $\widetilde{b}$-analogue of the b-ghost, it is pertinent to ask what is the generalised $\widetilde{c}$-ghost?

In analogy with the usual c-ghost, we shall look for conformal dimension -1 operator, satisfying the canonical relation

$$\left\{\oint \widetilde{b}, \widetilde{c}\right\} = 1.$$

This is not without its own complications. In particular, since $\widetilde{b}$ is at picture $+1$, $\widetilde{c}$ must be at picture -1 to satisfy the relation. It appears there is no suitable expression for $\widetilde{c}$ satisfying these conditions. However, since $\widetilde{L_{\alpha}}$ is on-shell, the picture-changing transformation is applicable to it. Since $L_1^\alpha = \{Q_0, \oint \widetilde{b}\}$ and picture-changing operators Γ and Γ^{-1} (direct and inverse) are BRST-invariant, one has $\Gamma^n L_1^\alpha = \{Q_0, \Gamma^n \oint \widetilde{b}\}$, so the picture-changing transform can be applied to generalised ghosts as well (even though they are off-shell).

It is convenient to bring $\oint \widetilde{b}$ to picture -1 by applying Γ^{-1} twice. The picture -1 expression for $\oint \widetilde{b}$ is given by:

$$\begin{aligned}\oint \widetilde{b}(w) = \oint \frac{dz}{2i\pi}\Bigg\{&-8\partial c c e^{3\phi-4\chi}\Bigg\{\frac{1}{2}P^{(2)}_{-\sigma}\Big[-\frac{3}{8}\partial^2 L + \frac{1}{4}\partial L P^{(1)}_{-16\phi+3\chi-3\sigma}\\
&+L\left(-\frac{3}{2}\partial^2\phi + \frac{11}{8}\partial^2\chi + \frac{3}{8}\partial^2\sigma - 4(\partial\phi)^2 + \frac{5}{8}(\partial\chi)^2 + \frac{1}{8}(\partial\sigma)^2\right.\\
&\left.+6\partial\phi\partial\chi - \frac{1}{2}\partial\phi\partial\sigma + \frac{7}{4}\partial\chi\partial\sigma\right)\Big] + \frac{1}{6}P^{(3)}_{-\sigma}\\
&\times\left(-\frac{3}{4}\partial L + L\left(\frac{1}{4}\partial\sigma - \frac{1}{2}\partial\phi\right)\right) + \frac{1}{48}P^{(4)}_{-\sigma}L\Bigg\}\\
&-ce^{2\chi-3\phi}\Bigg\{P^{(1)}_{-\sigma}\times\Big[-\frac{3}{8}\partial^2 F - \frac{1}{4}\partial F P^{(1)}_{\phi-2\chi+2\sigma} + F\\
&\times\left(\frac{1}{8}\partial^2\phi + \frac{15}{4}\partial^2\chi - \frac{1}{4}\partial^2\sigma + \frac{13}{8}(\partial\phi)^2\right.\\
&\left.-3(\partial\chi)^2 - \frac{5}{2}\partial\phi\partial\chi - \frac{3}{2}\partial\phi\partial\sigma + \partial\chi\partial\sigma\right)\Big]\\
&+\frac{1}{2}P^{(2)}_{-\sigma}\left(\frac{-1}{2}\partial F + F\left(-\frac{3}{2}\partial\phi - \partial\chi\right)\right) - \frac{1}{24}P^{(3)}_{-\sigma}\Bigg\}.\end{aligned}$$

Next, the conjugate $\widetilde{c}$-ghost, satisfying $\{\oint \widetilde{b}, \widetilde{c}\} = \Gamma$ (note that Γ is a picture-changing operator, i.e. picture-transformed unit operator 1) is given by:

$$\begin{aligned}\widetilde{c} = \frac{1}{2}e^{3\phi-\chi}&\left\{F\left(\frac{1}{3}P^{(3)}_{\phi-\chi} + \frac{1}{2}\partial\phi P^{(2)}_{\phi-\chi}\right) + GL\left(\frac{1}{2}P^{(2)}_{\phi-\chi} + \partial\phi P^{(1)}_{\phi-\chi} + \frac{1}{2}\partial F P^{(2)}_{\phi-\chi}\right)\right.\\
&\left.+\partial G L P^{(1)}_{2\phi-\chi} + G\partial L P^{(1)}_{\phi-\chi} + \frac{1}{2}\partial^2 G L + \partial G\partial L\right\}\\
&+be^{4\phi-2\chi}\left\{\frac{1}{2}GFP^{(1)}_{\phi-\chi-\frac{3}{4}\sigma}P^{(1)}_{\phi-\chi} + \frac{1}{12}LP^{(3)}_{\phi-\chi}P^{(1)}_{\phi-\chi-\frac{3}{4}\sigma}\right.\\
&\left.+\frac{1}{16}\partial L P^{(2)}_{\phi-\chi}P^{(1)}_{\phi-\chi-\frac{3}{4}\sigma}\right\}\\
&+\partial b b e^{5\phi-3\chi}\left\{-\frac{1}{8}P^{(1)}_{\phi-\chi-\frac{3}{4}\sigma}P^{(2)}_{2\phi-2\chi-\sigma} + \frac{1}{32}P^{(3)}_{2\phi-2\chi-\sigma}\right\}.\end{aligned}$$

Since the ground ring elements L_1^α and L_2^α can be shown to commute: $[L_1^\alpha, L_2^\alpha]=0$, the nilpotent BRST charge of ghost cohomology H_1 is by definition equal to

$$Q_1 = \widetilde{c}_1 L_1^\alpha + \widetilde{c}_2 L_2^\alpha \widetilde{c}_1 \equiv \widetilde{c}, \widetilde{c}_2 = \oint \widetilde{c}.$$

Computing the OPEs, it is straightforward to derive the manifest integrated expression for Q_1:

$$Q_1 = \oint \frac{dz}{2i\pi}\left[ce^{\phi}\left(GL + FP^{(1)}_{\phi-\chi}\right) + \frac{1}{4}e^{2\phi-\chi}\left(GF + \frac{1}{2}LP^{(2)}_{2\phi-2\chi-\sigma}\right) - \partial cc\xi L(z)\right].$$

This defines a *new BRST complex in RNS superstring theory*! It is an element of superconformal ghost cohomology H_1.

7.2 BRST charges of higher-order BRST cohomologies

In case of uncompactified critical RNS superstring theory, the α-symmetry is the only additional global space-time symmetry present in the theory. For RNS theories in non-critical dimensions or critical but compactified on S^1, there is a huge set of additional α-symmetries, due to interactions with the Liouville (or compactified) mode [24–26].

Thus, for a d-dimensional RNS theory, there exist $d+1$ additional α-symmetries of ghost cohomology H_1. Combined with $\frac{(d+1)(d+2)}{2}$ Poincaré symmetries (including the Liouville direction), these $d+2$ ghost–matter mixing symmetries of H_1 enlarge the space-time symmetry group from $SO(2,d)$ to $SO(2,d+1)$, bringing in the first extra dimension. Next, an H_2 cohomology can be shown to contain $(d+3)$ superconformal ghost number 2 α-symmetries which, when combined with Poincaré symmetries and α-symmetries of H_2, enlarge the space-time symmetry group to $SO(2,d+2)$, bringing in the second extra dimension. As an example of a typical α-generator of H_2:

$$L^\beta = \oint \frac{dz}{2i\pi} e^{2\phi} F(X,\psi) F(\varphi,\lambda)(z)$$

$$F(X,\psi) = \psi_m \partial^2 X^m - 2\partial\psi_m \partial X^m$$

$$F(\varphi,\lambda) = \lambda\partial^2\varphi - \partial\lambda\partial\varphi,$$

where ϕ and λ are the super-Liouville components (or those of a compactified direction). The ground rings of each $\widetilde{L}$ of H_2 consist of four elements, while the ring of each operator of H_3 is 6-dimensional. These rings are *non-commutative* and give rise to non-trivial symmetry algebras of local gauge symmetries:

$$L_i^\beta \equiv \partial_w^i L^\beta(w), \quad i=1,2,3,4$$
$$L_i^\gamma \equiv \partial_w^i L^\gamma(w), \quad i=1,2,3,4$$

satisfying $[L_i^I, L_j^J] = (m-n)C^{IJK} L_K^{i+j}$. So the BRST charges at higher levels are:

$$Q_2 = \sum_{j=1}^{3} \widetilde{c^{(2)}}_j L_j^\beta + f^{IJK}\frac{1}{2}c^{(2)}{}_i c^{(2)}{}_j \partial_{i+j}^2 \oint \widetilde{b}^{(2)}(w).$$

This construction can in principle be generalised to ghost cohomologies H_n of arbitrary n, with each cohomology rank having its own associate BRST charge Q_n; although we were only able to do it explicitly for $n \leq 3$.

Determining BRST cohomologies of Q_n for $n \geq 1$ is a challenging and interesting problem, although it looks plausible that each Q_n corresponds to RNS string theory with a certain background geometry.

7.3 Properties of Q_n: cohomologies

So far, we have been able to investigate the simplest non-trivial case $n=1$. The problem of investigating the higher-n cases is still to be addressed. In the critical uncompactified case the only non-trivial element of $Q_0 + Q_1$ is given by the massless gauge boson:

$$V(k) = \oint e^{-3\phi}\Big[(\vec{A}\cdot\partial\vec{X})(\vec{k}\cdot\partial\vec{X})(\vec{k}\cdot\vec{\psi}) \\ +(\vec{A}\cdot\vec{\psi})(\vec{k}\cdot\partial\vec{\psi})(\vec{k}\cdot\vec{\psi})(\vec{A}\cdot\vec{\psi})(\vec{k}\cdot\partial\vec{X})^2\Big]e^{i\vec{k}\cdot\vec{X}}$$

with $\vec{k}\cdot\vec{A}(\vec{k})=0$ and $(\vec{k})^2=0$. This operator is the element of H_{-3}. In the non-critical cases there are other massless modes. In particular, for $d=4$ there are seven extra massless vector bosons in the Q_0+Q_1 cohomology, altogether giving rise to an $SU(3)$ octet of gluons. *There are no non-trivial massive modes in the cohomology!* Such a field-theoretic behaviour is characterictic of string theories in AdS-type backgrounds, dual to CFTs.

7.4 New BRST charges and deformed pure spinors

In the pure spinor formalism of [27, 28], the standard BRST operator

$$Q_{\mathrm{PS}} = \oint \frac{dz}{2i\pi} \lambda^{\alpha} d_{\alpha}$$
$$d_{\alpha} = p_{\alpha} - \frac{1}{2}\gamma^{m}_{\alpha\beta}\partial X_m \theta^{\beta} - \frac{1}{8}(\theta\gamma^m\partial\theta)(\gamma_m\theta)_{\alpha}$$

(with $\alpha, \beta = 1, 2, \ldots, 16$) is nilpotent if $\lambda\gamma^m\lambda = 0$. This is called the *pure spinor condition*, and the OPE between two λ's is non-singular. The latter condition is ensured by the fact that in the standard pure spinor action λ is a free ghost field. This condition, however, can be relaxed with Q still remaining nilpotent even if OPE between pure spinors becomes singular, provided that the pure spinor constraint is still satisfied at a normal-ordered level and certain other constraints are fulfilled. For example, consider the most general OPE between $d_{\alpha}(z)d_{\beta}(w)$ around the midpoint:

$$d_{\alpha}(z)d_{\beta}(w) = -\frac{\gamma^{m}_{\alpha\beta}\prod^{(1)}_{m}\left(\frac{z+w}{2}\right)}{z-w} + (z-w)^0\gamma^{m_1\ldots m_3}_{\alpha\beta}\prod^{(2)}_{m_1\ldots m_3}\left(\frac{z+w}{2}\right)$$
$$+(z-w)\left\{\alpha_1\gamma^{m}_{\alpha\beta}\prod^{(3)}_{m} + \alpha_2\gamma^{m_1\ldots m_5}\prod^{(3)}_{m_1\ldots m_5}\left(\frac{z+w}{2}\right)\right\} \tag{7.10}$$

and suppose that λ satisfies the OPE

$$\lambda_{\alpha}(z)\lambda_{\beta}(w) \sim (z-w)^{-2}\gamma^{m}_{\alpha\beta}A_m\left(\frac{z+w}{2}\right) + O(z-w)$$

(no $(z-w)^0$ term means that the pure spinor constraint is fulfilled in a normal ordered (weaker) sense). Then the BRST charge is still nilpotent if either $\alpha_1 = 0$ or $A_m\Pi^{(3)}_m := 0$ (other singularities vanish upon evaluating traces of gamma-matrices). This is precisely the situation that is realised if one considers the RNS-PS map

$$\theta_{\alpha} = e^{\frac{1}{2}\phi}\Sigma_{\alpha}$$
$$\lambda_{\alpha} = \{Q_0, \theta_{\alpha}\} = -\frac{1}{4}be^{\frac{5}{2}\phi-2\chi}\Sigma_{\alpha} - \frac{1}{2}e^{\frac{3}{2}\phi-\chi}\gamma^{m}_{\alpha\beta}\partial X_m\tilde{\Sigma}^{\beta}$$
$$+ce^{\frac{1}{2}\phi}\left(\frac{1}{2}\Sigma_{\alpha}\partial\phi + \partial\Sigma_{\alpha}\right)$$

so that $\lambda_{\alpha}\lambda_{\beta} \sim \frac{1}{(z-w)^2}\partial bbe^{5\phi-4\chi}\gamma^{m}_{\alpha\beta}\psi_m + O(z-w)$.

It can be shown that, under such a RNS–PS identification, the pure spinor BRST charge is simply mapped to the RNS–BRST charge Q_0 (up to similarity transformation) and is therefore nilpotent [29]:

$$Q^{\mathrm{PS}} \to e^{-R} Q_0^{\mathrm{RNS}} e^{R}, \tag{7.11}$$

where $R = 32 \oint \frac{dz}{2i\pi} \partial c c e^{2\chi - 2\phi} \partial \chi(z)$. Thus the RNS theory is equivalent to a modified pure spinor theory with a double-pole singularity in the pure spinor OPE. This construction can be generalised to include the modified pure spinors with more singular OPE; remarkably, the RNS–BRST operators of higher ghost cohomologies are then reproduced, with the leading singularity order of pure spinor OPE related to the ghost cohomology order in RNS formalism. We have been able to show this for the $n = 1$ case and conjectured it for higher n's. Namely, in the $n = 1$ case one starts with $\widetilde{\theta^{\alpha}} = e^{\frac{3}{2}\phi} \Sigma^{\beta} \gamma^{m}_{\alpha\beta} (2\partial^2 X_m + \partial X_m \partial \phi)$ which is the dimension 0 primary field space-time spinor at ghost number $\frac{3}{2}$, *not* related to the previous ghost number $\frac{1}{2}$ version of θ by picture-changing. Next, one defines $\widetilde{\lambda}_{\alpha} = \{Q_0^{\mathrm{RNS}}, \widetilde{\theta}_{\alpha}\}$, keeping d^{α} unchanged at picture $-\frac{1}{2}$. A straightforward calculation gives $Q_1^{\mathrm{PS}} = \oint \widetilde{\lambda}_{\alpha} d^{\alpha} \to e^{-R} Q_1^{\mathrm{RNS}} e^{R}$ with the same R. This permits us to make the conjecture that *RNS superstring theory with the BRST operator Q_n of H_n is equivalent to modified pure spinor (PS) theory with singular pure spinor OPEs with the leading OPE singularity order given by* $6n^2 + 6n + 2$.

7.5 Conclusions

A hierarchy of surprising space-time α-symmetries in RNS superstring theories induces ground rings of matter–ghost mixing local gauge symmetries that can be classified in terms of ghost cohomologies H_n. Each ground ring induces the associated new BRST charge Q_n of H_n in RNS theory, corresponding to some deformed background geometry (AdS-type for $n = 1$) while each Q_n of RNS theory corresponds to deformed pure spinor superstring theory, with the leading OPE singularity order in the pure spinor formalism related to the ghost cohomology order of n in RNS formalism. This is a branch of superstring theory that appears to still be in its infancy and there remains much work to be done. Included among these future projects we count the investigation of cohomologies of Q_n for $n \geq 1$; the identification of geometries of underlying backgrounds and the associated building of string field theories around these backgrounds; and maybe even developing pure spinor-formulated string field theories inspired by the generalised RNS–PS map presented here.

References

[1] A. M. Polyakov, *Gauge Fields and Strings*, CHUR, Switzerland: Harwood (1987) 301 p. (Contemporary Concepts in Physics, Vol. 3).
[2] K. Wilson, *Phys. Rev.* **D10**: 2445 (1974).
[3] K. G. Wilson, *Nucl. Phys. Proc. Suppl.* **140**: 3 (2005) [arXiv:hep-lat/0412043].
[4] A. M. Polyakov, *Nucl. Phys. Proc. Suppl.* **68**, 1 (1998) [arXiv:hep-th/9711002].
[5] A. M. Polyakov, *Int. J. Mod. Phys. A* **14**: 645 (1999) [arXiv:hep-th/9809057].
[6] J. M. Maldacena, *Adv. Theor. Math. Phys.* **2**: 231 (1998).
[7] S. Gubser, I. Klebanov, and A. M. Polyakov, *Phys. Lett.* **B428**: 105–14.
[8] E. Witten, *Adv. Theor. Math. Phys.* **2**: 253–291 (1998).
[9] G. 't Hooft, *Nucl. Phys.* **B72**: 461 (1974).
[10] D. Berenstein, J. Maldacena, and H. Nastase, *JHEP* **0204**: 013 (2002) [hep-th/0202021].
[11] A. M. Polyakov and V. S. Rychkov, *Nucl. Phys.* **B594**: 272–86 (2001).
[12] J. E. Paton and H.-M. Chan, *Nucl. Phys.* **B10**: 516–20 (1969).
[13] E. Witten, *Nucl. Phys.* **B268**: 253 (1986).
[14] E. Witten, *Nucl. Phys.* **B276** (1986).
[15] D. Friedan, E. Martinec, and S. Shenker, *Nucl. Phys.* **B271**: 93 (1986).
[16] B. Zwiebach, *Nucl. Phys.* **B390**: 33–152 (1993) [hep-th/9206084].
[17] I. Y. Arefeva, P. B. Medvedev, and A. P. Zubarev, *Phys. Lett.* **B240**: 356 (1990).
[18] I. Y. Arefeva, P. B. Medvedev, and A. P. Zubarev, *Nucl. Phys.* **B341**: 464 (1990).
[19] N. Berkovits, *Nucl. Phys.* **B450**: 90 (1995) [hep-th/9503099].
[20] N. Berkovits, *Fortsch. Phys. (Progr. Phys.)* **48**: 31 (2000) [hep-th/9912121].
[21] C. Becchi, A. Rouet, and R. Stora, *Ann. Phys.* **98**: 287–321 (1976).
[22] A. Neveu and J. H. Schwarz, *Phys. Rev.* **D4**: 1109–11 (1971).
[23] D. Polyakov, arXiv:0906.3663, *IJMPA*, in press.
[24] D. Polyakov, arXiv:0905.4858, *IJMPA*, in press.
[25] D. Polyakov, *Int. J. Mod. Phys.* **A24**: 113–39 (2009).
[26] D. Polyakov, *Int. J. Mod. Phys.* **A22**: 2441 (2007).
[27] N. Berkovits, *JHEP* 0801:065 (2008).
[28] N. Berkovits, *JHEP* 0109:016 (2001).
[29] D. Polyakov, *Int. J. Mod. Phys.* **A24**: 2677–87 (2009).

Bibliography

Yu. M. Makeenko and A. A. Migdal, *Phys. Lett.* **B88**: 135 (1979).

8

Recent trends in superstring phenomenology

MASSIMO BIANCHI

We briefly review basic aspects of string theory and broadly discuss possible phenomenological scenario. We then focus on vacuum configurations with intersecting and/or magnetized unoriented D-branes. We show how a TeV-scale string tension may be compatible with the existence of large extra dimensions and how anomalous U(1)*s can give rise to interesting signatures at LHC or in cosmic rays. Finally, we discuss unoriented D-brane instantons as a source of non-perturbative effects that can contribute to moduli stabilization and SUSY breaking in combination with fluxes. We conclude with an outlook on holography and directions for future work.*

8.1 Foreword

More than 40 years after its original proposal, there is no experimental evidence for string theory or, else, a satisfactory – possibly holographic – description of the QCD string is not yet available.

Still, the Veneziano model predicts a massless vector boson in the open string spectrum and the Shapiro–Virasoro model a massless tensor boson in the closed string spectrum. These two particles can naturally be associated with the two best known forces in Nature: gravity and electromagnetism. After GSO projection, supersymmetry then guarantees the presence of massless fermions.

Moreover, string theory makes a definite – albeit incorrect – prediction for the number of space-time dimensions: 26 for bosonic strings, 10 for superstrings.

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

This basic fact led to many so-far unsuccessful attempts to get rid of the undesired extra dimensions. Calabi–Yau and orbifold compactifications, non-geometric Gepner models are the most famous examples.

Only relatively recently, it has been fully appreciated that one needs fluxes and non-perturbative effects for stabilizing moduli, breaking SUSY and eventually making contact with astroparticle phenomenology.

The plan of this chapter is to briefly review elementary features of string theory for non-experts[1] and to present the available phenomenological *scenario*. Starting from heterotic strings on CY 3-folds and unoriented strings (Type I and other Type II *un*-orientifolds), we will eventually discuss F-theory on elliptic CY 4-folds and M-theory on singular G_2-holonomy spaces and consider the role of fluxes, dualities and (Euclidean) branes in the characterization of an interesting portion of the landscape of vacua.

We will then focus on the case that most easily allows the embedding of the (N-MS)SM (not-necessarily minimal supersymmetric standard model) in string theory, i.e. intersecting and/or magnetized unoriented D-branes. We will show how a rather low string tension (TeV scale) is compatible with the existence of large extra dimensions and how anomalous $U(1)$s could give rise to spectacular signatures at LHC or in cosmic rays. Finally, we will discuss unoriented D-brane instantons as a source of non-perturbative superpotentials and other effects that crucially contribute to moduli stabilization and SUSY breaking in combination with fluxes.

We will conclude with an outlook and directions for future work.

8.2 String theory: another primer

General relativity is very successful in the description of gravity at large distances. However at very short distances, i.e. at energies comparable with the Planck mass $M_{\mathrm{Pl}} = (hc/G_{\mathrm{N}})^{1/2}$, the theory is inconsistent with quantum mechanics. Technically speaking, a quantum theory of gravity based on the Einstein–Hilbert action is not renormalizable since the Newton constant G_{N} – very much as the Fermi G_{F} for β decay – has dimension of an inverse mass squared. 'Miraculous' cancellations between boson and fermion loops, which make (rigid) supersymmetry the only viable candidate for the resolution of the hierarchy problem, persist at low orders in locally supersymmetric theories, aka (extended) supergravity. In particular, thanks to the relation of its amplitudes to those in $\mathcal{N} = 4$ super Yang–Mills (SYM) theory [3, 4], maximally extended $\mathcal{N} = 8$ supergravity in $D = 4$ seems to be UV finite beyond the order at which the onset of UV divergences was expected [5]. This has led to the conjecture that this 'unique' supergravity theory may be UV finite to all

[1] For a comprehensive exposition see, for example, [1, 2].

orders in perturbation theory and may accommodate non-perturbative stable states associated with regular charged black holes [6].

Notwithstanding this interesting possibility, in order to describe gravity at short distances, one may need a more radical change of perspective. Thanks to its very soft UV 'Regge behaviour', whereby amplitudes decay as $\mathcal{A} \approx e^{-\alpha' E^2}$ at large energies, string theory is the best available candidate for the unification of gravity with the other interactions in a consistent quantum theory. Perturbatively, there are two broad classes of string theories: those with only closed oriented strings and those with open and closed unoriented strings. The string spectrum is coded in the Regge trajectories which relate the mass M of a state to its spin S. For the closed-string states

$$\alpha' M_S^2 = 4(S-2), \tag{8.1}$$

which implies the existence of a massless symmetric tensor field ('graviton'). For the open-string states the relation becomes

$$\alpha' M_S^2 = (S-1), \tag{8.2}$$

and one may notice the presence of a massless vector boson ('photon'). In the early days of the hadronic string, in order to match the string states with the known resonances, the string tension $T = 1/2\pi\alpha'$ was taken to be of the order of the proton mass squared $T \approx M_{\mathrm{p}}^2 \approx (1\,\mathrm{GeV})^2$. In the unified string picture the string tension is supposed to be comparable with the Planck mass squared $T \approx M_{\mathrm{Pl}}^2 \approx (10^{19}\,\mathrm{GeV})^2$ so that the infinite tower of massive states decouples from the massless states at low energies, i.e. in the limit $\alpha' \to 0$. The very massive vibration modes with arbitrarily high spins, which at low energies can be neglected or better 'integrated out', may however play a crucial role in the early stages of evolution of the Universe. The exponential growth of the degeneracy of string states with energy signals a phase transition at the Hagedorn temperature $T_{\mathrm{H}} \approx 1/\sqrt{\alpha'}$. In recent times, the distinctive property of the string spectrum has been used to explain the microscopic origin of the Bekenstein–Hawking area formula for the macroscopic entropy of black holes in string theory [7].

In order to (first) quantize string theory and to display its symmetry principles, it is very convenient to proceed in analogy with Feynman's path-integral quantization of a point-particle. Instead of summing over 'world-lines', Alexander Polyakov proposed to sum over 'world-sheets', i.e. surfaces spanned by the string in its time evolution. The proper weight is the action of Brink, Di Vecchia and Howe (BDVH):

$$S[X,\gamma] = \frac{1}{4\pi\alpha'} \int_\Sigma G_{\mu\nu}(X)\partial_\alpha X^\mu \partial_\beta X^\nu \gamma^{\alpha\beta} \sqrt{|\gamma|}\, d^2\sigma, \tag{8.3}$$

where the string coordinates X^μ are maps from the two-dimensional world-sheet Σ with metric $\gamma_{\alpha\beta}$ into a D-dimensional 'target space' $\mathcal{M}_D$ with metric $G_{\mu\nu}$. In order to compute amplitudes of physical processes and get meaningful results, one has to divide out the infinite volume of the classical symmetry group. Indeed, the BDVH action is invariant under two-dimensional diffeomorphisms as well as under local Weyl rescalings of the world-sheet metric $\gamma_{\alpha\beta}$. Using the freedom of world-sheet coordinate changes, one can fix $\gamma_{\alpha\beta}$ to be conformally flat: $\gamma_{\alpha\beta} = e^{2\varphi}\delta_{\alpha\beta}$. At the classical level, the conformal factor φ decouples but at the quantum level, conformal anomalies are generated. Following the standard Faddeev–Popov (FP) procedure one finds that the conformal anomaly of the string coordinates X^μ is cancelled by the contribution of the FP ghosts b, c when $D = 26$, the critical dimension for the bosonic string. An analogous procedure shows that the critical dimension for the superstring, which requires the introduction of world-sheet fermionic partners Ψ^μ of the bosonic coordinates X^μ as well as the superconformal FP ghosts β, γ for the world-sheet gravitino χ_α, is $D = 10$. Thanks to the GSO (Gliozzi, Scherk and Olive) projection, the resulting superstring theory turns out to be supersymmetric and, as such, free of tachyons that plagued the bosonic strings.

Using conformal invariance, the Polyakov integral may be reduced to a perturbative series in the topology of Riemann surfaces. Once the functional integration over the world-sheet fields has been performed, one is left with a finite-dimensional integral over the Teichmüller parameters, which describe the 'shape' of Riemann surfaces with punctures corresponding to the insertions of vertex operators for the asymptotic external states. The measure of integration descends to the moduli space of Riemann surfaces if invariance under 'large' (disconnected) diffeomorphisms, generating the so-called mapping class group (aka modular group by abuse of terminology), is preserved at the quantum level. To summarize, conformal invariance and modular invariance are the guiding principles for the construction of (perturbatively) consistent closed-string models. For theories with open and unoriented strings, further subtle consistency conditions, such as tadpole cancellation and channel duality, are to be imposed.

The known supersymmetric string theories can all be formulated naturally in $D = 10$. According to the number of (Majorana–Weyl) supersymmetries, one has:

- Type I superstring – a theory of open and closed unoriented superstrings whose low-energy limit is chiral $\mathcal{N} = (1,0)$ supergravity coupled to $N = (1,0)$ SYM theory with gauge group $SO(32)$.
- Type IIA superstring – a theory of closed oriented superstrings, L–R asymmetric on the world-sheet, whose low-energy limit is the non-chiral $\mathcal{N} = (1,1)$ supergravity.

- Type IIB superstring – a theory of closed oriented superstrings, L–R symmetric on the world-sheet, whose low-energy limit is the chiral $\mathcal{N} = (2,0)$ supergravity.
- Heterotic string – a L–R asymmetric combination of bosonic and fermionic strings whose low-energy limit is chiral $\mathcal{N} = (1,0)$ supergravity coupled to $\mathcal{N} = (1,0)$ SYM theory with gauge group $Spin(32)/Z_2$ or $E(8) \times E(8)$.

The first and the last of these theories, though chiral, are anomaly free thanks to the Green–Schwarz mechanism which involves the antisymmetric tensor field $B_{\mu\nu} = -B_{\nu\mu}$, present in the massless spectrum. The Type IIB theory, though chiral, is anomaly-free thanks to a remarkable cancellation among the contributions of two left-handed gravitini, two right-handed dilatini and a 4-form potential $C^{(+)}_{\mu\nu\rho\sigma}$ whose field strength has to be self-dual.

All the above theories contain a metric $G_{\mu\nu}$, an antisymmetric tensor $B_{\mu\nu}$ and a scalar (dilaton) Φ in their massless bosonic spectra. In addition, in the open-string spectrum of the Type I theory and in the Neveu–Schwarz (NS) spectrum of the heterotic theory, massless vector bosons A^a_μ appear. In Type II theories the massless NS–NS spectrum is accompanied by a rich spectrum of tensor fields coming from the massless Ramond–Ramond (R–R) sector. The Type IIA contains an abelian vector field C_μ and a 3-form potential $C_{\mu\nu\rho}$, while the Type IIB theory, apart from the above $C^{(+)}_{\mu\nu\rho\sigma}$, contains another antisymmetric tensor $C_{\mu\nu}$ and a second dilaton C. The vacuum expectation value of the dilaton field $\langle\Phi\rangle$, which is undetermined in string perturbation theory, plays the role of string coupling constant g_s, i.e. of string-loop expansion parameter.

8.2.1 Green–Schwarz mechanism and heterotic strings

25 years ago,[2] Michael Green and John Schwarz proposed a mechanism for anomaly cancellation in Type I superstrings (unoriented open and closed strings) with $SO(32)$ gauge group [8]. Assigning non-linear gauge transformation properties to the R–R 2-form

$$\delta_{\mathrm{YM}} C_{\mu\nu} = Tr(\alpha_{\mathrm{YM}} F_{\mu\nu}) + \ldots$$

and including a local counterterm in the effective action of the form

$$L_{\mathrm{GS}} = C \wedge Tr(F \wedge F \wedge F \wedge F) + \ldots$$

restores gauge invariance at the quantum level. The Green–Schwarz mechanism works for $E(8) \times E(8)$. The milestone paper [8] left open the problem of how

2 Anomaly cancellations in supersymmetric D = 10 gauge theory and superstring theory, by Michael B. Green and John H. Schwarz, *Physics Letters B*, Volume 149, Issues 1–3, 13 December 1984, pages 117–122. Received 10 September 1984.

to embed this gauge group in string theory. Indeed, the abstract to [8] ended with the statement '... A superstring theory for $E(8) \times E(8)$ has not yet been constructed'. The answer to the problem is the *heterotic string*[3] that combines the left-moving superstrings with right-moving bosonic strings [9]. The difference $(26-10=16)$ between critical dimensions accounts for the rank of either $SO(32)$ or $E(8) \times E(8)$.

Compactifications on Calabi–Yau 3-folds were soon recognized as a very promising framework to produce chiral $\mathcal{N}=1$ supersymmetric models with $E(6)$ grand unified group, via 'standard embedding' of the spin connection in one of the two $E(8)$ groups. The net number of generations of chiral fermions turns out to be solely determined by the topology of the CY space, i.e. $N_{\text{gen}} = |\chi(\text{CY})|/2$, where $\chi(\text{CY}) = 2(h_{11} - h_{21})$ is the Euler characteristic [10]. For roughly 10 years, from 1985 to 1995, heterotic model building attracted the attention of the vast majority of string theorists in the hope that non-standard embeddings ('stable holomorphic vector bundles'), giving rise to $SO(10), SU(5), ..., SU(3) \times SU(2) \times U(1)$ gauge groups and $N_{\text{gen}} \neq \chi(\text{CY})/2$, could produce acceptable particle phenomenology [11].

Many problems, however, were left open [2]:

- CY compactifications introduce many new undetermined parameters, associated with deformations of the complex and Kähler structures. These may be associated to vacuum expectation values of the so-called *moduli* fields, i.e. massless scalars which are flat directions of the scalar potential, and be interpreted as internal components of the metric, antisymmetric tensors and gauge fields.
- It is hard to get precisely $N_{\text{gen}} = 3$ and no vector-like pairs, and to avoid chiral exotics and/or fractional charges.
- Barring large threshold corrections, there is a significant discrepancy between $M_{\text{Pl}} \approx M_s \sqrt{\hat{V}_{\text{int}}/g_s}$ and $M_{\text{GUT}} \approx M_{\text{comp}} \approx M_s/\hat{V}_{\text{int}}^{1/6}$, since $\alpha_{\text{GUT}} \approx g_s^2/\hat{V}_{\text{int}} \approx 1/25$.
- Without turning on fluxes and/or including non-perturbative effects, there is no viable mechanism for SUSY breaking that avoids tachyons and produces a very small dark-energy density.

Still, there were important achievements emerging from the perturbative studies of superstrings [2]:

- Enormous advances in the understanding of world-sheet conformal invariance and its implications on Riemann surfaces were achieved.

[3] Heterotic string by David J. Gross, Jeffrey A. Harvey, Emil Martinec and Ryan Rohm, *Physics Review Letters* Volume 54, 1985, pages 502–505. Received 21 November 1984.

- A large yet restricted class of on-shell low-energy effective supergravity actions were derived from superstring scattering amplitudes.
- The very extended nature of the fundamental constituents leads to the equivalence between large and small volumes known as T-duality. Its action exchanges windings and 'standard' KK momenta for closed strings as well as Neumann b.c. and Dirichlet b.c. for open strings.
- Mirror symmetry exchanging complex and Kähler deformations were 'experimentally' found and then systematized in the so-called quantum/stringy geometry.
- Non-geometric models (orbifolds, free fermions, Gepner, ...) were proposed that reproduce interesting features of CY compactifications without the need for sophisticated algebrogeometric constructions.

8.2.2 Type I and other Type II un-orientifolds

As already mentioned, theories with unoriented open and closed strings have the virtue of predicting gravitational interactions mediated by closed strings and gauge interactions mediated by open strings.

In 1987, Augusto Sagnotti proposed viewing the Type I superstring in $D=10$ as a *parameter space orbifold* or *open string descendant* of the Type IIB superstring [12].

For almost 10 years, from 1986 to 1996, Type I model building *without* direct mention of D-branes and Ω-planes was systematically developed by the group in Tor Vergata [13]. The results were rather surprising:

- In addition to Type I strings, non-supersymmetric (non)-tachyonic models were identified in $D \leq 10$ [14].
- In a peculiar asymmetric super-ghost picture, vertex operators for R–R states were proposed that involve the potentials C_{p+1} rather than their field-strengths F_{p+2} [17], thus pre-dating Polchinski's observation on the minimal coupling of D-branes to R–R fields.
- In toroidal compactifications, the role of Wilson lines, T-dual to D-brane separations, in the adjoint breaking of the gauge group and of a non-zero but quantized $B^{\mathrm{NS-NS}}$ in its rank reduction was recognized [17].
- Consistent chiral models with $\mathcal{N}=(1,0)$ SUSY in $D=6$ were constructed that contain several tensor multiplets [15] and lead to a generalized GS mechanism [16].
- The first instance of an $\mathcal{N}=1$ SUSY chiral model in $D=4$ with $N_{\mathrm{gen}}=3$ was obtained after compactification on the T^6/Z_3 orbifold [18].
- The role of magnetic fluxes that support chiral fermions and trigger (partial) SUSY breaking was emphasized [19].

- A very detailed understanding of boundary conformal field theory and its implications was built up [20].

It is fair to say, however, that the geometric picture whereby open strings terminate on D-branes [21] which carry R–R charge [22] was crucial in all later developments, including holography.

8.2.3 Branes and dualities

Polchinski's observation [22] that D-branes are hypersurfaces where open strings can end and are sources of R–R charges was not isolated. In addition to fundamental strings ('F1-branes'), which are charged with respect to the NS–NS 2-form $B^{(2)\mathrm{NS-NS}}_{\mu\nu}$, string theory indeed admits other 1/2 BPS extended solitonic solutions, called p-branes [23]. Various p-branes couple to different higher-rank antisymmetric tensor fields and extremize the relation between their tension and charge density $T_p = |Q_p|$. Depending on the $(p+1)$-forms the p-brane couples to, one has:

- NS5-brane couples to $\widetilde{B}^{\mathrm{NS-NS}}_6$, which is the magnetic dual to $B^{\mathrm{NS-NS}}_2$ that couples to fundamental strings [24].
- D-branes that couple to $C^{(p+1)\mathrm{R-R}}_{\mu_1\mu_2\ldots\mu_{p+1}}$
  - D0-, D2-, D4-, D6-, D8-branes are BPS in Type IIA,
  - D(-1)-, D1-, D3-, D5-, D7-, D9-branes are BPS in Type IIB.

The generalized 'tension', i.e. energy per unit world-volume, is a distinctive feature of each kind of brane: for fundamental strings $T_{\mathrm{F1}} \approx 1/\alpha'$, for NS5-branes $T_{\mathrm{NS5}} \approx 1/g_s^2(\alpha')^3$, for Dp-branes $T_{\mathrm{Dp}} \approx 1/g_s(\alpha')^{\frac{p+1}{2}}$.

Moreover, although one knows very well the microscopic world-sheet description of a single fundamental string and, to a lesser extent, a single NS5-brane, it is not clear how to describe the dynamics of a stack of strings let alone for NS5-branes. On the contrary, at least at low energy, the dynamics of a stack of N Dp-branes is known to be governed by the dimensional reduction to $D = p+1$ of $D = 10$ $\mathcal{N} = (1,0)$ SYM in $D = 10$, with $U(N)$ gauge group [25].

The list of extended objects includes the rather peculiar orientifold planes [21], or simply Ω-planes, which play a crucial role in the construction of theories with unoriented strings. Ω-planes are (perturbatively) non-dynamical BPS solitons with negative tension, which can neutralize the R–R charge of D-branes. Ω-planes act as 'mirrors' both on the target space and on the world-sheet: $X_\perp(z,\bar{z}) \to -X_\perp(\bar{z},z)$, and allow D-branes with $SO(N)$ and $Sp(2N)$ gauge groups or (anti)symmetric tensor representations of $U(N)$.

8.2.4 String duality conjectures

As a result of the existence of BPS p-branes, which may become light under suitable conditions, different string theories may offer complementary descriptions of the same physics in different regimes [26, 27].

The best-tested duality conjectures are:

- Type IIA at strong coupling exposes one more dimension and gives rise to supergravity in $D=11$ at low energy [26, 27]. The dilaton is related to the diagonal component of the metric in the extra dimension $\phi \approx G_{11,11}$, while the R–R vector and the NS–NS 2-form arise from the mixed components of the metric and 3-form, respectively. D0-branes (D-particles) are nothing but KK excitations while all other p-branes in Type IIA[4] can be related to M2-branes, sourced by the 3-form, their magnetic dual M5-branes in $D=11$ and purely gravitational solitons.
- $SO(32)$ heterotic and Type I strings are S-dual in $D=10$ [28–30]. In the string 'frame', their effective actions coincide up to the field redefinitions

$$\phi_{\mathrm{H}} = -\phi_{\mathrm{I}}, \quad G^{\mathrm{H}}_{\mathrm{MN}} = e^{-\phi_{\mathrm{I}}} G^{\mathrm{I}}_{\mathrm{MN}}. \tag{8.4}$$

  The $SO(32)$ heterotic string can be identified with the Type I D-string since $T_{\mathrm{H}} = T_{\mathrm{I}}/g_s^{\mathrm{I}}$, while Type I fundamental strings being unoriented are not BPS objects.[5] The duality entails strong $\leftrightarrow$ weak coupling exchange, i.e. $g_s^{\mathrm{I}} = 1/g_s^{\mathrm{H}}$. The duality has not only been accurately tested at the level of BPS saturated couplings in toroidal and some orbifold compactifications [31], but also at the level of some massive (non)-BPS states such as the 128 spinorial heterotic states that correspond to a Type I D-particle in $D=10$ [32].
- The heterotic string on T^4 is dual to Type IIA on $K3$ [33]. In this duality the heterotic string should be viewed as the Type IIA NS5-branes wrapped around a $K3$ surface and vice versa. The duality has been accurately tested at the level of BPS saturated couplings in $D=6$ and in toroidal compactifications.

The picture that emerges is that all known superstring theories can be related to one another by the action of some discrete transformation of the so-called U-duality group [26]. The theory that subtends the network has been (provisionally) called M-theory [27]. U-duality transformations are the discrete remnants of the non-compact continuous symmetries of (extended) supergravities. The existence of (discrete) U-duality transformations acting non-linearly on massless scalar fields opens the door to new kinds of compactifications or rather vacuum configurations, whereby scalars

4 Except for D8-branes which are domain walls in $D=10$ separating regions with different Romans' mass $m=F_0$, which is not allowed in $D=11$.

5 They are 'sourced' by the NS–NS 2-form which cannot fluctuate in Type I, being projected out by Ω.

are non-trivially identified under monodromies. The most prominent one, as we will see later, goes under the name of F-theory and exploits the $SL(2)$ symmetry of Type IIB superstrings acting non-linearly on the complexified dilaton $\tau = \chi + ie^{-\phi}$ [34]. In the F-theory picture, τ can be considered as the complex structure modulus of an auxiliary torus (elliptic curve) that is non-trivially fibred over 'physical' space-time.

Quite astonishingly, the study of brane dynamics led Juan Maldacena to conjecture another quite remarkable duality, i.e. the holographic correspondence between Type IIB in five-dimensional anti De Sitter space and $\mathcal{N} = 4$ SYM on its four-dimensional (conformal) boundary [35]. For the first time the long-sought-for string description of large-N gauge theories was made concrete, albeit for a very special gauge theory that is superconformal and does not 'confine' in the usual sense. Moreover, the holographic principle for gravity of 't Hooft and Susskind finds a precise realization in Maldacena correspondence, that proved to be a gold mine. In addition to being tested and generalized to many other forseeable contexts, including warped extra-dimensional scenarios [36], it has led to integrability and to unforseeable applications to non-relativistic condensed matter and astrophysical systems [37].

Progress in holography at a microscopic world-sheet level has so far been hampered by the long-standing problem of quantizing strings in the presence of R–R fluxes. Even in maximally symmetric cases such as $AdS_5 \times S^5$, little is known. Some timid steps can be taken by resorting to the pure spinor formalism [50] or other hybrids [81], but the path is long. Except for very special cases such as 'deformed' (super) WZW models [52] or some plane waves [53], not even the free string spectrum is known, let alone the interactions. In a related line of investigation, the recent discovery that $\mathcal{N} = 4$ SYM theory may be integrable in the planar limit has opened the way to significant progress in the calculation of anomalous dimensions of gauge-invariant composite operators holographically dual to masses of string states in AdS with R–R flux [54]. Although explicit computations even for the simplest operators, belonging in the so-called Konishi multiplet [55], may require days of computer time [56], one should try to be optimistic!

8.3 Phenomenological scenarios

In any serious attempt to embed the standard model in a fundamental theory, two key features should be taken into account: the chiral spectrum of elementary fermions and the existence of a light scalar field, the Higgs boson. In order to accommodate chirality and protect the hierarchy between the weak scale (100 GeV) and any fundamental scale (10^{16-19} GeV), vacuum configurations with (softly broken) $\mathcal{N} = 1$ SUSY in $D = 4$ are definitely the best candidate. In string theory, this leads to the following options which we will momentarily review without entering into technical details:

- Heterotic strings on CY 3-folds with non-trivial vacuum gauge bundles.
- Intersecting and/or magnetized D-branes in Type I strings and other Type II *un*-orientifolds.
- F-theory on elliptic CY 4-folds.
- M-theory on singular G_2 holonomy 7-spaces.

In order to stabilize moduli and break SUSY, one needs to turn on internal fluxes [38], i.e. gauge some of the continuous symmetries present in the low-energy supergravity description [39], and include non-perturbative effects from Euclidean p-branes wrapping internal cycles, i.e. 'p-brane instantons' [40].

8.3.1 Heterotic on CY 3-folds with fluxes/branes

For historical reasons, the first class of phenomenologically viable models correspond to heterotic strings on CY 3-folds. Before including fluxes and (non)-perturbative effects, given a suitable CY 3-fold, one has to find a non-trivial vacuum gauge bundle to produce some SUSY extension of the standard model.

One has to rely heavily on the algebrogeometric construction of stable holomorphic vector bundles [1]. The choice of the structure group determines the resulting GUT group: $SU(3), SU(4), SU(5)$ respectively correspond to $E(6), SO(10), SU(5)$ GUTs from the visible $E(8)$. The hidden $E(8)$ gives rise to gaugino condensation that can produce a field-dependent superpotential $W = \Lambda^3_{\text{SYM}}(\phi_i)$. In addition to the obvious dependence on the 4-d dilaton, gauge couplings and $\Lambda = Me^{8\pi^2/\beta g^2(M)}$ indeed acquire one-loop threshold corrections that depend on the compactification moduli. This simple W leads to a runaway potential. A variant with several gaugino condensates at different scales, known as 'race-track' models, seems more appealing [41]. A very interesting possibility is to have 'local Grand Unification' [42]. Yet the inclusion of 3-form and metric (torsion) fluxes as well as branes seems unavoidable for moduli stabilization and SUSY breaking [43].

NS–NS fluxes are amenable to a world-sheet analysis and may even lead to interesting non-geometric constructions involving T-duality. The resulting T-folds correspond to gaugings of discrete non-compact symmetries and cannot be interpreted as geometric manifolds, in that the gluing of the various patches requires identifications which are not simply diffeomorphisms. This is only possible for theories with stable extended solitons (p-branes) and goes beyond the point-particle description. A trick that allows us to deal with some non-geometric construction in more geometric terms at least for tori is the 'doubling' of all internal coordinates [44].

It is, however, an unsolved problem how to include (wrapped) NS5s even in geometric backgrounds, let alone in non-geometric configurations. On the contrary, D-branes admit a neat CFT description, that will prove extremely useful in the sequel, and can easily be adapted to non-geometric backgrounds [45].

8.3.2 Unoriented D-brane worlds

As pioneered by Augusto Sagnotti and the Tor Vergata group [13], vacuum configurations with open and closed unoriented strings can be thought of as 'descendants' of left–right symmetric theories with oriented strings. In more geometric terms, this means including D-branes and, for consistency, Ω-planes [46].

Depending on the choice of parent theory and Ω projection, one has the following options:

- Type I, i.e. Type IIB with Ω9-plane completely wrapping the internal space and implementing a Z_2 projection of the closed-string spectrum under world-sheet parity Ω. Additional Ω5-planes wrapping (collapsed) 2-cycles may be present. One can then include (magnetized) D9-branes and D5-branes wrapping holomorphic 2-cycles with F_3 fluxes
- Type IIB with Ω3-planes localized at points in the internal space and implementing a Z_2 projection of the closed-string spectrum $\Omega_B = \Omega(-)^{F_L}\mathcal{I}$ where $(-)^{F_L}$ denotes L-mover fermion parity and $\mathcal{I}$ is a holomorphic involution, acting locally as $z^I \leftrightarrow -z^I$, whose fixed-point set is a collection of points and, possibly, 4-cycles, the latter leading to Ω7-planes. One can then include D3-branes and (magnetized) D7-branes wrapping holomorphic 4-cycles with $F_3 + \tau H_3 = G_3$ fluxes.
- Type IIA with Ω6-planes wrapping a 3-cycle and implementing a Z_2 projection of the closed string spectrum $\Omega_A = \Omega\widetilde{\mathcal{I}}$, where Ω denotes world-sheet parity and $\widetilde{\mathcal{I}}$ is an antiholomorphic involution (locally $z^I \leftrightarrow -\bar{z}^I$) whose fixed-point set is a collection of Ω6-planes. One can then include intersecting D6-branes wrapping SLAGs (special Lagrangian 3-cycles) with various F_0, F_2, F_4 R–R fluxes as well as some H_3 and metric fluxes.

In any case, Ω-planes with negative tension and R–R charge are needed for tadpole cancellation (Gauss law), i.e. anomaly cancellation (G-S strikes back!) [47,48], e.g.

$$N_{D3} + \frac{1}{8\pi^3}\int_{\mathcal{M}_6} H_3 \wedge F_3 + \frac{1}{8\pi^2}\int_{\mathcal{W}_4^{(7)}} Tr(F_2 \wedge F_2) = \frac{1}{4}N_{\Omega 3}$$

for Type IIB *un*-orientifolds with D3- and D7-branes. In some cases, 'orientifold' singularities can be resolved (non)-perturbatively, as for instance in F-theory, whereby $\Omega 7$-planes are replaced by collections of mutually non-local 7-branes [49]. Since T-duality or appropriate generalizations, such as mirror symmetry, should relate the above *un*-orientifolds to one another, the very notion of Ω-planes which is intrinsically perturbative and quite counter-intuitive, in-so-far as their negative tension is concerned for instance, should be replaced by more sophisticated constructions generalizing F-theory by the extension of $SL(2)$ to larger monodromy groups. In the same vein, geometric and non-geometric fluxes should be unified in a common framework and associated with different choices of embedding tensor that encodes the allowed gaugings in the supergravity description [39].

8.3.3 M-theory phenomenology

Type IIA at strong coupling grows another dimension [27]: $\mathcal{M}_{10} \to \mathcal{M}'_{11} \approx \mathcal{M}_{10} \times S^1$. The dilaton, parameterizing the extra circle, is fibred over the base. Non-trivial fibrations correspond to intrinsically non-perturbative constructions (in g_s). In particular, compactifications on CY 3-folds are replaced by compactifications on 7-manifolds. In order to preserve $\mathcal{N}=1$ SUSY in $D=4$ (4 out of the original 32 supercharges), the holonomy should be G_2 [57]. Indeed, under $SO(7) \to G_2$ the spinor decomposes according to $\mathbf{8} \to \mathbf{1}+\mathbf{7}$. The singlet corresponds to unbroken SUSY. If the manifold is smooth, however, no chirality is generated. One has to consider M-theory on singular G_2 holonomy spaces. Very much as in CY 3-folds, manifolds with G_2 holonomy admit a globally defined 3-form, Φ_3, known as the associative 3-form. The co-associative 4-form $*\Phi_3 = \widetilde{\Phi}_4$, represents a natural choice of G_4 flux. In the G(2)–MSSM framework proposed in [58], however, no G_4 flux is needed. Built-in non-perturbative effects such as 'racetrack'-like gaugino condensates [41] produce a superpotential that can stabilize all moduli and produce peculiar (N–MS)SM spectra. The explicit construction of the singular manifold of G_2 holonomy with the desired properties, such as the appropriate singularity structure and a small number of moduli to start with, is still an open issue.

At any rate, despite the inclusion of a class of non-perturbative effects, the M-theory framework is only computationally reliable in the low-energy supergravity approximation since a truly microscopic description is still lacking. Only very recently, significant progress has been made in the understanding of the dynamics of stacks of M2-branes [59]. Very little is known about coincident M5-branes.

8.3.4 F-theory on elliptic CY 4-folds

Another promising class of compactifications goes under the name of F-theory [34]. In more mundane terms, F-theory describes Type IIB configurations with varying dilaton $\tau = \tau(z)$, where z are the 'internal' coordinates. In particular, in the presence of mutually non-local 7-branes, the dilaton varies and undergoes $SL(2, Z)$ monodromy when carried around the location of the 7-branes.

In a Weierstrass model for the auxiliary torus fibration: $y^2 = x^3 + f_8(z)x + g_{12}(z)$, with z-coordinates on a complex 3-fold base $\mathcal{B}$, the locations of 7-branes are coded in the discriminant locus

$$\Delta_{24}(z) = 27 f_8(z)^3 + 4 g_{12}^2(z) = \prod_i (z - z_i). \tag{8.5}$$

The local behaviour of the complexified dilaton is

$$2\pi i \tau(z) \approx \log(z - z_i). \tag{8.6}$$

In the orientifold limit, in which mutually non-local 7-branes form a bound state with monodromy $\tau \to \tau - 4$, the dilaton is locally constant (up to monodromy) and the rest of the 7-branes are D7-branes [49].

Recently, there has been a noteworthy revival of local GUTs from F-theory on CY 4-folds. In local models with 7-branes wrapping Del Pezzo surfaces [60], it has been shown that one can turn on 'magnetic' F_{Y} fluxes breaking $SU(5)$ to the SM. A specific choice of four stacks $\mathcal{S}, \mathcal{P}, \mathcal{Q}, \mathcal{R}$ seems capable of producing all known (chiral) matter at intersections, as well as the necessary Yukawa's at triple intersections, with a concrete prediction for the Cabibbo angle, or rather the texture of the CKM matrix. For consistency, one has to assume that the local model could be embedded in a global setting in such a way that additional branes and fluxes could stabilize the moduli controlling the strength of the Yukawa couplings [61].

Notwithstanding its elegance in reproducing low-energy phenomenology, F-theory lacks even more than M-theory a concrete microscopic description. Most of the results are valid only at low energy in that they are based on complex algebraic properties that can only control massless or light states. Massive strings and branes are completely beyond reach, together with all the resulting 'threshold' corrections to gauge and Yukawa couplings.

From now on, we will focus on the perturbative framework of open and closed unoriented strings, that are in principle amenable to a world-sheet analysis at any energy scale.

8.4 Intersecting vs magnetized branes

In theories with open strings, it is important to observe that magnetized D-branes support massless chiral fermions [19, 62]. The chiral asymmetry is governed by an index theorem [63]

$$\mathcal{I}_{ab} = \frac{W_a W_b}{2\pi^n n!} \int_{\mathcal{M}_{2n}} (q_a F_a + q_b F_b)^n,$$

where W denotes the wrapping number and F the magnetic field on the two stacks of branes where the two ends of an open string with charges $q = \pm 1$ are located. Magnetized branes are characterized by the modified boundary conditions for the string coordinates:

$$\partial_n X^\mu = 2\pi\alpha' F^\mu{}_\nu \partial_t X^\nu. \tag{8.7}$$

Under T-duality $\partial_t Y \leftrightarrow \partial_n Y$, yielding the boundary conditions for a rotated brane [64]. Thus, relatively magnetized branes become intersecting branes under T-duality. The intersection angles are related to the magnetic fluxes via

$$\tan(\beta_{ab}^{ij}) = q_a F_a^{ij} + q_b F_b^{ij}.$$

The SUSY condition for branes at an angle on $T^6 \approx T^2_{12} \times T^2_{34} \times T^2_{56}$ is

$$\beta_{ab}^{12} + \beta_{ab}^{34} + \beta_{ab}^{56} = 0.$$

More generally, one can consider ‘co-isotropic’ branes, i.e. magnetized branes intersecting with other (magnetized) branes.

A peculiar feature of this class of models, i.e. vacuum configurations with intersecting and/or magnetized D-branes, is the lack of a GUT structure. Since the (tree-level) gauge couplings [46]

$$(g_{\rm YM}^{(a)})^2 = \frac{g_s}{\sqrt{\det(\mathcal{G}_{(a)} + 2\pi\alpha'\mathcal{F}_{(a)})}}$$

depend on the magnetic flux $\mathcal{F}_{(a)}$ and metric $\mathcal{G}_{(a)}$ induced on the internal cycle wrapped by a given stack of D-branes, GUT would be quite unnatural or unnecessary.

8.4.1 Low string tension and large extra dimensions

Very large internal volume is compatible with low string tension ($M_s \approx$ TeV) and $g_s << 1$ [65] since for Type I strings and unoriented relatives

$$M_{\rm Pl} \approx \frac{M_s\sqrt{V_{\rm int}}}{g_s}.$$

One arrives at the idea of unoriented D-brane worlds, whereby (N–MS)SM lives on a stack of branes while gravity lives in the 10-d bulk. The weakness of gravity arises from ‘dilution’ in the extra dimensions

$$M^2_{\rm Pl,(4)} = M^{D-2}_{\rm Pl,(D)} L^{D-4}_{\rm int}.$$

A slightly alternative scenario relies on 'warped' extra dimensions [36]. In the simplest instance the universe is a slice of *AdS* space: gravity is localized on the UV brane at one end, with large warp factor, while (N–MS)SM is localized on the TeV brane at the other end, with exponentially small warp factor.

Spectacular signals may be detected at LHC from KK excitations and/or Regge recurrences! Indeed, many string amplitudes, such as MHV amplitudes and SUSY-related ones, relevant for collisions at accelerators, contain a simple Veneziano form-factor [66–70]

$$\mathcal{A}^{\mathrm{MHV}}_{4,\mathrm{ST}}(s,t,u) = \mathcal{A}^{\mathrm{MHV}}_{4,\mathrm{FT}}(s,t,u)\frac{\Gamma(1-\alpha' s)\Gamma(1-\alpha' t)}{\Gamma(1+\alpha' u)}.$$

In order to embed the SM in this context, one needs at least four stacks of intersecting and/or magnetized branes [2, 46]:

$$U(3)\times U(2)\times U(1)\times U(1)' \to SU(3)_C\times SU(2)_W\times U(1)_Y\times U(1)^3.$$

Hypercharge embedding $Q_{\mathrm{Y}} = aQ_{\mathrm{A}} + bQ_{\mathrm{B}} + cQ_{\mathrm{C}} + dQ_{\mathrm{D}}$ leads to two favourite choices [71, 72] and many other possibilities [73].

The presence of extra anomalous $U(1)$s requires not only a generalized GS mechanism but also generalized Chern–Simons couplings [74]. Some non-anomalous $U(1)$s become massive anyway due to anomalies in higher dimensions [75]. Baryon and lepton numbers may appear as 'global' remnants of some (anomalous) gauge symmetries.

8.4.2 Anomalous U(1)s and generalized CS couplings

The need for GCS couplings can easily be argued for. Chiral anomaly receives a one-loop contribution from the fermion determinant

$$\delta\mathcal{L}_{1-\mathrm{loop}} = t_{ijk}\alpha^i F^j\wedge F^k,$$

with $t_{ijk} = \sum_f Q^i_f Q^j_f Q^k_f$ totally symmetric and the (algebraic) sum running over chiral fermions (+ left, − right).

Closed-string axions β^I in (twisted) R–R sectors couple to open-string vector bosons A^i_μ and behave as Stückelberg fields

$$\mathcal{L}_{ax} = (d\beta^I - M^I_i A^i)\wedge *(d\beta^I - M^I_i A^i) + C^I_{jk}\beta_I F^j\wedge F^k.$$

The axionic shifts can be gauged by (anomalous) vector bosons

$$\delta A^i = d\alpha^i \qquad \delta\beta^I = M^I_i\alpha^i.$$

As a result, one finds

$$\delta\mathcal{L}_{ax} = M_i^I \alpha^i C_{jk}^I \beta_I F^j \wedge F^k$$

with $M_i^I C_{jk}^I$ not totally symmetric. Introducing generalized CS couplings

$$\mathcal{L}_{\text{GCS}} = E_{ij,k} A^i \wedge A^j \wedge F^k$$

so that $\delta\mathcal{L}_{\text{GCS}} = E_{ij,k}\alpha^i \wedge F^j \wedge F^k$ leads to complete anomaly cancellation [74]

$$t_{ijk} - M_k^I \alpha^i C_{ij}^I = -E_{ij,k}.$$

GCS terms can involve the vector boson that gauges hypercharge and some vector boson Z' that gauges an (anomalous) $U(1)$. After electroweak symmetry breaking, this would lead to a trilinear coupling of Z' with Z and γ with peculiar signals at LHC for $M_{Z'}$ at the TeV scale.

At LHC one could detect the decay $Z' \to Z + \gamma$, by observing a rather sharp line in the spectrum of γ. For instance, in the Z' rest frame one has

$$E_\gamma = \frac{M_{Z'}^2 - M_Z^2}{2M_{Z'}}.$$

Similarly, one can observe other decay channels, such as $Z' \to Z + Z^*$ or $Z' \to \gamma + \gamma^*$, that require one (slightly) off-shell product, for the decay amplitude not to vanish due to Bose statistics.[6]

Even if it were visible at LHC, the above processes would not necessarily be a signature of string theory but could be taken as circumstantial evidence for low-tension/LED strings.

Moreover, in this class of models, the axino, the fermionic superpartner of the axio-dilaton contributing to anomaly cancellation, may turn out to be the LSP (rather the 'standard' neutralino) and thus a candidate for dark matter. Analysis of signals from cosmic ray experiments, such as PAMELA and FERMI, are under way [76–80].

8.5 Unoriented D-brane instantons

Despite some success in embedding (N–MS)SM in vacuum configurations with open and unoriented strings, there are some problems with interactions at the perturbative level:

- In $U(5)$ (SUSY) GUTs, the Yukawa couplings $H^d_{\mathbf{5}^*_{-1}} F^c_{\mathbf{5}^*_{-1}} A_{\mathbf{10}_{+2}}$ are allowed but $H^u_{\mathbf{5}_{+1}} A_{\mathbf{10}_{+2}} A_{\mathbf{10}_{+2}}$ are forbidden by (global, anomalous) $U(1)$ invariance. Though

[6] For massive vector boson pairs, it may be possible to bypass the hypotheses of the Landau–Yang theorem and get a non-vanishing amplitude even for on-shell products. We thank A. Racioppi for private communications on this point.

compatible with $SU(5)$ symmetry, there is no way to generate the necessary ε^{abcde} in perturbation theory via Chan–Paton factors.

- R-handed (s)neutrino masses $W_{\rm M} = M_{\rm R} N N$ are forbidden by e.g. $U(1)_{\rm B-L}$ in Pati–Salam-like models $SO(6) \times SO(4) \rightarrow SU(3) \times SU(2)_{\rm L} \times SU(2)_{\rm R} \times U(1)_{\rm B-L}$.
- The μ-term in MSSM $W_\mu = \mu H_1 H_2$ is typically forbidden by extra (anomalous) $U(1)$s or their 'global' remnants.

All the above couplings can, however, be generated by 'stringy' instantons, in particular unoriented D-brane instantons [40], that result from wrapping Euclidean branes around non-contractible cycles in the internal space [81–83].

There are two broad classes of unoriented D-brane instantons:

- 'Gauge' instantons correspond to EDp-branes wrapping the same cycle C as a stack of background D(p+4)-branes, they induce $F = *_4 F$ and generate effects with strength

$$e^{-W_{p+1}(C)/g_s \ell_s^{p+1}} = e^{-1/g^2_{\rm YM}}.$$

The resulting system has roughly speaking 4 N–D directions (space-time) and allows us to make the (super)ADHM construction rather intuitive!

- 'Exotic' instantons correspond to EDp′-branes wrapping a cycle C' not wrapped by any stack of background D(p+4)-branes. In some cases $F \wedge F = *_8 F \wedge F$ [84], the strength is given by

$$e^{-W_{p'+1}(C')/g_s \ell_s^{p'+1}} \neq e^{-1/g^2_{\rm YM}}$$

and roughly speaking the system has 8 N–D directions (space-time + internal).

8.5.1 Non-perturbative effects

For phenomenological purposes it is crucial to identify which non-perturbative effects can be induced by the two classes of unoriented D-brane instantons.

'Gauge instantons' may generate VY–ADS-like superpotentials of the form

$$W = \frac{\Lambda^\beta}{\phi^{\beta-3}} \quad \text{with} \quad \Lambda = M_s e^{-\frac{8\pi^2}{\beta g^2_{YM}(M_s)}}$$

that decay for large VEVs $\langle\phi\rangle$, where the effective gauge theory is asymtotically free. The rate is governed by β, the one-loop β-function coefficient of the (quiver) gauge theory living on the background stack of D-branes.

'Exotic instantons' may generate non-perturbative superpotentials of the form

$$W = M_s^{3-n} e^{-S_{EDp'}(C')} \phi^n \quad (n = 1, 2, 3, ...)$$

that can grow for large VEVs $\langle\phi\rangle$.

In both cases, the thumb rule for a non-vanishing result is the presence of two exact fermionic zero-modes not lifted by (D-term) interactions. Explicit computations for intersecting branes and branes at singularities (quiver gauge theories) have confirmed the above pattern [85–89].

(Partial) moduli stabilization and SUSY breaking in combination with (R–R) fluxes and field-dependent FI terms for anomalous $U(1)$s can be achieved in some special cases [90–92]. Unoriented D-brane instantons with a different structure of fermionic zero-modes may lead to threshold corrections to gauge couplings and other higher-derivative F-terms. One can learn a lot about multi-instantons or poly-instantons exploiting heterotic/Type I duality [93–95].

8.5.2 Fluxes, dualities and ... landscape

In the intersecting brane picture it is relatively easy to identify (non)-SUSY flux vacua with $\Lambda \leq 0$. Indeed, many near-horizon geometries of this kind of configurations are of the form $AdS_4 \times G/H$ [96] and a variant of the attractor mechanism [97] is at work. Most of the would-be moduli are stabilized by the flux-generated scalar potential. Unfortunately, very much as for $AdS_5 \times S^5$, one has some control only in the low-energy supergravity approximation. The spectrum of massive strings is completely unknown, as is customary in the presence of (R–R) fluxes. Conditions for decoupling of massive KK excitations can be met in some cases, but deriving their spectrum by means of harmonic analysis on G/H may be rather involved and tedious.

In addition to these technical problems, the crucial issue is the uplift to De Sitter with a very small cosmological constant $0 < \Lambda \ll M$ [98, 99]. All the mechanisms proposed so far have some drawbacks. Moreover, there seems to be some tension between moduli stabilization and chirality to the extent that, at present, there is not a single model enjoying at least two of the following three properties:

- embed N–MSSM with its chiral content;
- stabilize all the moduli;
- break supersymmetry with very small dark energy.

Maybe one should not look for specific (classes of) models but rather extract statistical predictions and correlations among macroscopic observables in the vast landscape of vacua, e.g. number of generations vs rank of the gauge group, Yukawa couplings vs average gauge couplings, ...

8.6 Outlook

Another possibility that recently emerged is what may be termed 'holographic phenomenology'. It may well be possible that, thanks to the holographic principle, string theory could find some experimental evidence in previously unimagined systems, such as:

- quark–gluon plasma in HIC [100];
- superconductors [101];
- hydrodynamics [102];
- Kerr BHs and superstars [103];
- gravity at a Lifshitz point [104].

In conclusion, there are several viable phenomenological scenarios in string theory but no one is to be particularly preferred since predictions are still rather vague.

To a certain extent, vacuum configurations with open and unoriented strings that allow for low string tension and large extra dimensions can be falsified even at LHC. They tend to predict extra (anomalous) Z', Regge recurrences and KK excitations at the TeV scale. In some cases, a non-standard 'neutralino', the axino, may be the LSP and thus the relevant WIMP candidate for dark matter.

A complete global picture is still lacking: the large scale structure of the string landscape is to be explored at a much deeper level with an eye wide open onto holography phenomenology. String theory owes much to *1984* [8, 9], but it is not *doublethink* [105].

Acknowledgments

I would like to thank P. Anastasopoulos, C. Angelantonj, E. Kiritsis, S. Kovacs, J. F. Morales, G. Pradisi, G. C. Rossi, A. Sagnotti, Ya. Stanev for long-lasting and very enjoyable collaboration on various aspects of string phenomenology and M. Samsonyan and L. Lopez for suggestions on the manuscript. I would like to thank STIAS and NIThep for hospitality during completion of this project and the organizers of *Foundations of Space and Time – Reflections on Quantum Gravity* for creating a stimulating atmosphere. This work was partially supported by the ERC Advanced Grant n.226455 *'Superfields'* and by the Italian MIUR-PRIN contract 20075ATT78 *'Symmetries of the Universe and of the fundamental interactions'*.

References

[1] M. B. Green, J. H. Schwarz and E. Witten, *Superlisting Theory: I, II*, Cambridge Monographs on Mathematical Physics. Cambridge, UK: Cambridge University Press (1987); 596 pp.

[2] E. Kiritsis, *String Theory in a Nutshell*. Princeton, NJ: Princeton University Press (2007); 588 pp.
[3] Z. Bern, L. J. Dixon, M. Perelstein, D. C. Dunbar and J. S. Rozowsky, *Class. Quant. Grav.* **17**, 979 (2000) [arXiv:hep-th/9911194].
[4] M. Bianchi, H. Elvang and D. Z. Freedman, JHEP **0809**, 063 (2008) [arXiv:0805.0757 [hep-th]].
[5] Z. Bern, J. J. Carrasco, L. J. Dixon, H. Johansson and R. Roiban, *Phys. Rev. Lett.* **103**, 081301 (2009) [arXiv:0905.2326 [hep-th]].
[6] M. Bianchi, S. Ferrara and R. Kallosh, arXiv:0910.3674 [hep-th].
[7] A. Strominger and C. Vafa, *Phys. Lett. B* **379**, 99 (1996) [arXiv:hep-th/9601029].
[8] M. B. Green and J. H. Schwarz, *Phys. Lett. B* **149**, 117 (1984).
[9] D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, *Phys. Rev. Lett.* **54**, 502 (1985).
[10] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, *Nucl. Phys. B* **258**, 46 (1985).
[11] L. E. Ibanez, J. Mas, H. P. Nilles and F. Quevedo, *Nucl. Phys. B* **301**, 157 (1988).
[12] A. Sagnotti (1987) [arXiv:hep-th/0208020].
[13] For a review see e.g. C. Angelantonj and A. Sagnotti, *Phys. Rept.* **371**, 1 (2002) [Erratum ibid. **376**, 339 (2003)] [arXiv:hep-th/0204089].
[14] M. Bianchi and A. Sagnotti, *Phys. Lett. B* **211**, 407 (1988). G. Pradisi and A. Sagnotti, *Phys. Lett. B* **216**, 59 (1989). M. Bianchi and A. Sagnotti, *Phys. Lett. B* **231**, 389 (1989).
[15] M. Bianchi and A. Sagnotti, *Phys. Lett. B* **247**, 517 (1990). M. Bianchi and A. Sagnotti, *Nucl. Phys. B* **361**, 519 (1991). C. Angelantonj, M. Bianchi, G. Pradisi, A. Sagnotti and Y. S. Stanev, *Phys. Lett. B* **387**, 743 (1996) [arXiv:hep-th/9607229].
[16] A. Sagnotti, *Phys. Lett. B* **294**, 196 (1992) [arXiv:hep-th/9210127].
[17] M. Bianchi, G. Pradisi and A. Sagnotti, *Nucl. Phys. B* **376**, 365 (1992).
[18] C. Angelantonj, M. Bianchi, G. Pradisi, A. Sagnotti and Y. S. Stanev, *Phys. Lett. B* **385**, 96 (1996) [arXiv:hep-th/9606169].
[19] M. Bianchi and Y. S. Stanev, *Nucl. Phys. B* **523**, 193 (1998) [arXiv:hep-th/9711069]. C. Angelantonj, I. Antoniadis, G. D'Appollonio, E. Dudas and A. Sagnotti, *Nucl. Phys. B* **572**, 36 (2000) [arXiv:hep-th/9911081].
[20] M. Bianchi, G. Pradisi and A. Sagnotti, *Phys. Lett. B* **273**, 389 (1991). G. Pradisi, A. Sagnotti and Y. S. Stanev, *Phys. Lett. B* **381**, 97 (1996) [arXiv:hep-th/9603097].
[21] J. Dai, R. G. Leigh and J. Polchinski, *Mod. Phys. Lett. A* **4**, 2073 (1989).
[22] J. Polchinski, *Phys. Rev. Lett.* **75**, 4724 (1995) [arXiv:hep-th/9510017].
[23] P. K. Townsend, arXiv:hep-th/9507048.
[24] A. Dabholkar, J. P. Gauntlett, J. A. Harvey and D. Waldram, *Nucl. Phys. B* **474**, 85 (1996) [arXiv:hep-th/9511053]. D. S. Berman, *Phys. Rept.* **456**, 89 (2008) [arXiv:0710.1707 [hep-th]].
[25] E. Witten, *Nucl. Phys. B* **460**, 335 (1996) [arXiv:hep-th/9510135].
[26] C.M. Hull and P.K. Townsend, *Nucl. Phys. B* **438**, 109 (1995) [arXiv:hep-th/9410167].
[27] E. Witten, *Nucl. Phys. B* **443**, 85 (1995) [arXiv:hep-th/9503124].
[28] J. Polchinski and E. Witten, *Nucl. Phys. B* **460**, 525 (1996) [arXiv:hep-th/9510169].
[29] C. M. Hull, *Phys. Lett. B* **357**, 545 (1995) [arXiv:hep-th/9506194].
[30] A. Dabholkar, *Phys. Lett. B* **357**, 307 (1995) [arXiv:hep-th/9506160].
[31] C. Bachas, C. Fabre, E. Kiritsis, N. A. Obers and P. Vanhove, *Nucl. Phys. B* **509**, 33 (1998) [arXiv:hep-th/9707126]. E. Gava, J. F. Morales, K. S. Narain and G. Thompson, *Nucl. Phys. B* **528**, 95 (1998) [arXiv:hep-th/9801128]. M. Bianchi, E. Gava, F. Morales and K. S. Narain, *Nucl. Phys. B* **547**, 96 (1999) [arXiv:hep-th/9811013].

[32] M. Frau, L. Gallot, A. Lerda and P. Strigazzi, *Nucl. Phys. B* **564**, 60 (2000) [arXiv:hep-th/9903123].
[33] E. Kiritsis, N. A. Obers and B. Pioline, JHEP **0001**, 029 (2000) [arXiv:hep-th/0001083].
[34] C. Vafa, *Nucl. Phys. B* **469**, 403 (1996) [arXiv:hep-th/9602022].
[35] J. M. Maldacena, *Adv. Theor. Math. Phys.* **2**, 231 (1998) [*Int. J. Theor. Phys.* **38**, 1113 (1999)] [arXiv:hep-th/9711200]. E. Witten, *Adv. Theor. Math. Phys.* **2**, 253 (1998) [arXiv:hep-th/9802150]. S. S. Gubser, I. R. Klebanov and A. M. Polyakov, *Phys. Lett. B* **428**, 105 (1998) [arXiv:hep-th/9802109].
[36] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 4690 (1999) [arXiv:hep-th/9906064]. L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999) [arXiv:hep-ph/9905221].
[37] U. Gursoy, E. Kiritsis, L. Mazzanti and F. Nitti, *Nucl. Phys. B* **820**, 148 (2009) [arXiv:0903.2859 [hep-th]].
[38] S. B. Giddings, S. Kachru and J. Polchinski, *Phys. Rev. D* **66**, 106006 (2002) [arXiv:hep-th/0105097].
[39] G. Dall'Agata and S. Ferrara, *Nucl. Phys. B* **717**, 223 (2005) [arXiv:hep-th/0502066]. B. de Wit, H. Nicolai and H. Samtleben, JHEP **0802**, 044 (2008) [arXiv:0801.1294 [hep-th]]. M. Grana, J. Louis, A. Sim and D. Waldram, JHEP **0907**, 104 (2009) [arXiv:0904.2333 [hep-th]].
[40] For a recent review see e.g. R. Blumenhagen, M. Cvetic, S. Kachru and T. Weigand, arXiv:0902.3251 [hep-th].
[41] L. J. Dixon, 'Supersymmetry breaking in string theory', 15th APS Division. of Particles and Fields General Meeting, Houston, TX, Jan 3–6, 1990. DPF Conf. 1990: 811–822 (QCD161:A6:1990:V.2).
[42] H. P. Nilles, S. Ramos-Sanchez and P. K. S. Vaudrevange, arXiv:0909.3948 [hep-th].
[43] A. Strominger, *Nucl. Phys. B* **274**, 253 (1986).
[44] C. M. Hull and R. A. Reid-Edwards, arXiv:hep-th/0503114.
[45] M. Bianchi, *Nucl. Phys. B* **805**, 168 (2008) [arXiv:0805.3276 [hep-th]].
[46] R. Blumenhagen, B. Kors, D. Lust and S. Stieberger, *Phys. Rept.* **445**, 1 (2007) [arXiv:hep-th/0610327]. A. M. Uranga, JHEP **0901**, 048 (2009) [arXiv:0808.2918 [hep-th]].
[47] M. Bianchi and J. F. Morales, JHEP **0003**, 030 (2000) [arXiv:hep-th/0002149].
[48] G. Aldazabal, D. Badagnani, L. E. Ibanez and A. M. Uranga, JHEP **9906**, 031 (1999) [arXiv:hep-th/9904071].
[49] A. Sen, *Phys. Rev. D* **55**, 7345 (1997) [arXiv:hep-th/9702165].
[50] N. Berkovits, arXiv:hep-th/0209059.
[51] N. Berkovits and B. C. Vallilo, *Nucl. Phys. B* **624**, 45 (2002) [arXiv:hep-th/0110168].
[52] C. Candu, V. Mitev, T. Quella, H. Saleur and V. Schomerus, arXiv:0908.0878 [hep-th].
[53] D. E. Berenstein, J. M. Maldacena and H. S. Nastase, *AIP Conf. Proc.* **646**, 3 (2003).
[54] G. Arutyunov and S. Frolov, *J. Phys. A* **42**, 254003 (2009) [arXiv:0901.4937 [hep-th]].
[55] M. Bianchi, S. Kovacs, G. Rossi and Y. S. Stanev, JHEP **0105**, 042 (2001) [arXiv:hep-th/0104016].
[56] N. Gromov, V. Kazakov and P. Vieira, arXiv:0906.4240 [hep-th].
[57] B. S. Acharya and E. Witten, arXiv:hep-th/0109152.
[58] B. S. Acharya, K. Bobkov, G. L. Kane, J. Shao and P. Kumar, *Phys. Rev. D* **78**, 065038 (2008) [arXiv:0801.0478 [hep-ph]].
[59] J. Bagger and N. Lambert, *Phys. Rev. D* **75**, 045020 (2007) [arXiv:hep-th/0611108]. A. Gustavsson, *Nucl. Phys. B* **811**, 66 (2009) [arXiv:0709.1260 [hep-th]]. O. Aharony, O. Bergman, D. L. Jafferis and J. Maldacena, JHEP **0810**, 091 (2008) [arXiv:0806.1218 [hep-th]].

[60] J. J. Heckman, A. Tavanfar and C. Vafa, arXiv:0906.0581 [hep-th].

[61] R. Blumenhagen, T. W. Grimm, B. Jurke and T. Weigand, arXiv:0908.1784 [hep-th].

[62] A. Abouelsaood, C. G. Callan, C. R. Nappi and S. A. Yost, *Nucl. Phys. B* **280**, 599 (1987). C. Bachas, arXiv:hep-th/9503030. I. Antoniadis and T. Maillard, *Nucl. Phys. B* **716**, 3 (2005) [arXiv:hep-th/0412008].

[63] M. Bianchi and E. Trevigne, JHEP **0601**, 092 (2006) [arXiv:hep-th/0506080]. M. Bianchi and E. Trevigne, JHEP **0508**, 034 (2005) [arXiv:hep-th/0502147]. I. Antoniadis, A. Kumar and T. Maillard, *Nucl. Phys. B* **767**, 139 (2007) [arXiv:hep-th/0610246].

[64] V. Balasubramanian and R. G. Leigh, *Phys. Rev. D* **55**, 6415 (1997) [arXiv:hep-th/9611165].

[65] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, *Phys. Lett. B* **436**, 257 (1998) [arXiv:hep-ph/9804398].

[66] S. Cullen, M. Perelstein and M. E. Peskin, *Phys. Rev. D* **62**, 055012 (2000) [arXiv:hep-ph/0001166].

[67] E. Dudas and C. Timirgaziu, *Nucl. Phys. B* **716**, 65 (2005) [arXiv:hep-th/0502085].

[68] D. Chialva, R. Iengo and J. G. Russo, *Phys. Rev. D* **71**, 106009 (2005) [arXiv:hep-ph/0503125].

[69] M. Bianchi and A. V. Santini, *JHEP* **0612**, 010 (2006) [arXiv:hep-th/0607224].

[70] L. A. Anchordoqui, H. Goldberg, D. Lust, S. Nawata, S. Stieberger and T. R. Taylor, *Nucl. Phys. B* **821**, 181 (2009) [arXiv:0904.3547 [hep-ph]].

[71] I. Antoniadis, E. Kiritsis, J. Rizos and T. N. Tomaras, *Nucl. Phys. B* **660**, 81 (2003) [arXiv:hep-th/0210263].

[72] G. Aldazabal, S. Franco, L. E. Ibanez, R. Rabadan and A. M. Uranga, *JHEP* **0102**, 047 (2001) [arXiv:hep-ph/0011132].

[73] P. Anastasopoulos, T. P. T. Dijkstra, E. Kiritsis and A. N. Schellekens, *Nucl. Phys. B* **759**, 83 (2006) [arXiv:hep-th/0605226].

[74] P. Anastasopoulos, M. Bianchi, E. Dudas and E. Kiritsis, *JHEP* **0611**, 057 (2006) [arXiv:hep-th/0605225].

[75] P. Anastasopoulos, *JHEP* **0308**, 005 (2003) [arXiv:hep-th/0306042].

[76] P. Anastasopoulos, F. Fucito, A. Lionetto, G. Pradisi, A. Racioppi and Y. S. Stanev, *Phys. Rev. D* **78**, 085014 (2008) [arXiv:0804.1156 [hep-th]]. F. Fucito, A. Lionetto, A. Mammarella and A. Racioppi, arXiv:0811.1953 [hep-ph].

[77] R. Armillis, C. Coriano', M. Guzzi and S. Morelli, *Nucl. Phys. B* **814**, 15679 (2009) [arXiv:0809.3772 [hep-ph]].

[78] E. Dudas, Y. Mambrini, S. Pokorski, A. Romagnoni and M. Trapletti, *JHEP* **0903**, 011 (2009) [arXiv:0809.5064 [hep-th]].

[79] P. Anastasopoulos, E. Kiritsis and A. Lionetto, *JHEP* **0908**, 026 (2009) [arXiv:0905.3044 [hep-th]].

[80] F. Fucito, A. Lionetto, A. Mammarella and A. Racioppi, arXiv:0811.1953 [hep-ph].

[81] M. Dine, N. Seiberg, X. G. Wen and E. Witten, *Nucl. Phys. B* **278**, 769 (1986).

[82] K. Becker, M. Becker and A. Strominger, *Nucl. Phys. B* **456**, 130 (1995) [arXiv:hep-th/9507158].

[83] E. Witten, *Nucl. Phys. B* **460**, 541 (1996) [arXiv:hep-th/9511030]. M. R. Douglas, arXiv:hep-th/9512077.

[84] M. Billo, L. Ferro, M. Frau, L. Gallot, A. Lerda and I. Pesando, *JHEP* **0907**, 092 (2009) [arXiv:0905.4586 [hep-th]].

[85] M. Billo, M. Frau, I. Pesando, F. Fucito, A. Lerda and A. Liccardo, *JHEP* **0302**, 045 (2003) [arXiv:hep-th/0211250]. M. Billo', L. Ferro, M. Frau, F. Fucito, A. Lerda and J. F. Morales, *JHEP* **0810**, 112 (2008) [arXiv:0807.1666 [hep-th]].

[86] R. Blumenhagen, M. Cvetic and T. Weigand, *Nucl. Phys. B* **771**, 113 (2007) [arXiv:hep-th/0609191]. M. Cvetic, R. 2. Richter and T. Weigand, *AIP Conf. Proc.* **957**, 30 (2007).
[87] L. E. Ibanez and A. M. Uranga, *JHEP* **0703**, 052 (2007) [arXiv:hep-th/0609213]. L. E. Ibanez, A. N. Schellekens and A. M. Uranga, *JHEP* **0706**, 011 (2007) [arXiv:0704.1079 [hep-th]].
[88] M. Bianchi and E. Kiritsis, *Nucl. Phys. B* **782**, 26 (2007) [arXiv:hep-th/0702015]. M. Bianchi, F. Fucito and J. F. Morales, *JHEP* **0707**, 038 (2007) [arXiv:0704.0784 [hep-th]].
[89] B. Florea, S. Kachru, J. McGreevy and N. Saulina, *JHEP* **0705**, 024 (2007) [arXiv:hep-th/0610003].
[90] M. Bianchi, F. Fucito and J. F. Morales, *JHEP* **0908**, 040 (2009) [arXiv:0904.2156 [hep-th]].
[91] I. Garcia-Etxebarria, F. Saad and A. M. Uranga, *JHEP* **0705**, 047 (2007) [arXiv:0704.0166 [hep-th]].
[92] R. Argurio, M. Bertolini, S. F. Franco and S. Kachru, *Fortsch. Phys.* **55**, 644 (2007).
[93] R. Blumenhagen and M. Schmidt-Sommerfeld, *JHEP* **0807**, 027 (2008) [arXiv:0803.1562 [hep-th]].
[94] M. Bianchi and J. F. Morales, *JHEP* **0802**, 073 (2008) [arXiv:0712.1895 [hep-th]].
[95] P. G. Camara, E. Dudas, T. Maillard and G. Pradisi, *Nucl. Phys. B* **795**, 453 (2008) [arXiv:0710.3080 [hep-th]].
[96] C. Caviezel, P. Koerber, S. Kors, D. Lust, D. Tsimpis and M. Zagermann, *Class. Quant. Grav.* **26**, 025014 (2009) [arXiv:0806.3458 [hep-th]].
[97] S. Ferrara, R. Kallosh and A. Strominger, *Phys. Rev. D* **52**, 5412 (1995) [arXiv:hep-th/9508072].
[98] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, *Phys. Rev. D* **68**, 046005 (2003) [arXiv:hep-th/0301240].
[99] J. P. Conlon, C. H. Kom, K. Suruliz, B. C. Allanach and F. Quevedo, *JHEP* **0708**, 061 (2007) [arXiv:0704.3403 [hep-ph]].
[100] S. S. Gubser and A. Karch, arXiv:0901.0935 [hep-th].
[101] S. A. Hartnoll, arXiv:0903.3246 [hep-th].
[102] S. Bhattacharyya, S. Minwalla and S. R. Wadia, *JHEP* **0908**, 059 (2009) [arXiv:0810.1545 [hep-th]].
[103] T. Hartman, W. Song and A. Strominger, arXiv:0908.3909 [hep-th]. J. de Boer, K. Papadodimas and E. Verlinde, arXiv:0907.2695 [hep-th].
[104] P. Horava, *Phys. Rev. D* **79**, 084008 (2009) [arXiv:0901.3775 [hep-th]].
[105] G. Orwell, *1984*, First Edition (Secker & Warburg, 1949).

9

Emergent spacetime

ROBERT DE MELLO KOCH & JEFF MURUGAN

We give an introductory account of the AdS/CFT correspondence in the $\frac{1}{2}$-BPS sector of $\mathcal{N}=4$ super Yang–Mills theory. Six of the dimensions of the string theory are emergent in the Yang–Mills theory. In this chapter we suggest how these dimensions and local physics in these dimensions emerge. The discussion is aimed at non-experts.

9.1 Introduction

The problem of quantizing gravity has proved to be a difficult one. To solve this problem, it seems to be necessary to answer the question "What is spacetime?" This challenges the most basic assumptions we are used to making; a radical new idea may be needed. Further, the hope of any guidance from experiment seems to be out of the question. One might conclude that the situation is hopeless. Drawing on recent insights from the AdS/CFT correspondence, we are nonetheless, optimistic.

The AdS/CFT correspondence [1] claims an exact equality between $\mathcal{N}=4$ super Yang–Mills theory in flat $(3+1)$-dimensional Minkowski spacetime and Type IIB string theory on an asymptotically $\mathrm{AdS}_5\times \mathrm{S}^5$ background. Type IIB string theory is a theory of closed strings; at least within string perturbation theory, theories of closed strings provide a consistent UV completion of gravity. The fact that such an equality exists is highly unexpected and non-trivial, and (as we will try to convince the reader) can be used to gain insight into the nature of spacetime. George Ellis opened the *Foundations of Space and Time* workshop by holding up two fingers and asking "are there an infinite or a finite number of places a particle could occupy between my fingers?" We don't know the answer to George's question. However, we

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

hope to convince the reader that the AdS/CFT correspondence provides a detailed and concrete framework within which this question can be tackled.

We know how to formulate $\mathcal{N}=4$ super Yang–Mills theory as a path integral. We do not yet understand how to formulate quantum gravity in asymptotically $\mathrm{AdS}_5\times\mathrm{S}^5$ backgrounds. It seems somewhat natural then, to use the $\mathcal{N}=4$ super Yang–Mills theory as a definition for quantum gravity in asymptotically $\mathrm{AdS}_5\times\mathrm{S}^5$ backgrounds. The puzzle then becomes one of translating and interpreting the quantum field theory, as a quantum theory of gravity. Conceptually this is challenging and we do not have any simple arguments that would explain why a higher-dimensional gravity theory is encoded in the dynamics of a quantum field theory. Technically, it's tough too. The relation between the radius of the AdS space (measured in units of the string length l_s) and the 't Hooft coupling[1] ($\lambda=g_{\mathrm{YM}}^2N$)

$$\frac{R_{\mathrm{AdS}}^4}{l_s^4}=\lambda$$

shows that in the limit of small curvatures (where we could have hoped to recognize a familiar description of geometry) the field theory is strongly coupled and hence we do not know how to do the relevant field theory calculations. Conversely, if we compute things perturbatively in the field theory, we are studying the small λ limit where curvature corrections are important and our usual notions of geometry are probably not useful.

Fortunately, there is a way to proceed. Thanks to the large amount of supersymmetry enjoyed by the theory, there do exist quantities that are protected from corrections. If one chooses carefully, these quantities can be computed at weak coupling and the result can then be extrapolated to strong coupling. The most protected states of the theory, preserving half of the maximal amount of supersymmetry, are called the $\frac{1}{2}$-BPS sector. This is the laboratory in which we will work.

In Section 9.2 we will give some arguments for the simplicity of the $\frac{1}{2}$-BPS sector. In Section 9.3 we will explain how the dictionary between the gauge theory and the gravity theory is organized – it's organized according to the $\mathcal{R}$-charge[2] of the operators of the field theory. Section 9.4 introduces a set of variables, the Schur polynomials, which provide a beautiful organization of the degrees of freedom of the theory. In Sections 9.5, 9.6 and 9.7 we explain how to describe gravitons, strings

[1] Recall that by suitably rescaling the fields one can arrange things so that all g_{YM}^2 dependence factors out as an overall $\frac{1}{g_{\mathrm{YM}}^2}$ factor in front of the action. It is then clear that g_{YM}^2 plays the role of $\hbar$ for the quantum field theory.

[2] The $\mathcal{R}$-charge is a conserved charge associated specifically with supersymmetric theories. Recall that an internal symmetry is one whose generators commute with all of the spacetime symmetry generators. An $\mathcal{R}$ symmetry is one whose generators commute with all of the bosonic spacetime symmetries but fail to commute with the fermionic supercharges.

and branes of the string theory using the field theory language, and in Section 9.8 we explain how new backgrounds (the so-called LLM geometries) arise. Section 9.9 is reserved for discussion.

There are a number of papers that have significantly influenced our point of view and have had an impact on our research. Among these we mention [2–9].

9.2 Simplicity of the $\frac{1}{2}$-BPS sector

We study $\mathcal{N}=4$ super Yang–Mills theory on $R\times S^3$. The field content of $\mathcal{N}=4$ super Yang–Mills theory includes six Hermitian scalars transforming in the adjoint of the gauge group. We group the six real scalars into three complex fields as follows:

$$Z=\phi_1+i\phi_2,\qquad Y=\phi_3+i\phi_4,\qquad X=\phi_5+i\phi_6\,.$$

The $\frac{1}{2}$-BPS chiral primary operators we focus on can be built from a single complex combination (we use Z in what follows). Using a total of n Zs, there is a distinct operator for each partition of n. Given the partition with parts $\{n_i\}$, the corresponding operator is $\prod_i \mathrm{Tr}(Z^{n_i})$. There is a one-to-one correspondence between these operators and $\frac{1}{2}$-BPS representations of $\mathcal{R}$-charge n [3].

A beautiful argument, due to Berenstein [4], demonstrates the simplicity of the $\frac{1}{2}$-BPS sector.[3] Consider a time slicing of $\mathrm{AdS}_5\times\mathrm{S}^5$, which gives the Hamiltonian

$$H=\frac{(\Delta-J)+\varepsilon\Delta}{\varepsilon}\,,$$

where Δ is the dilatation operator and J is the $\mathcal{R}$-charge under which Z has one unit of charge. In the limit $\varepsilon\to 0$, any state with $\Delta-J>0$ will have a huge energy and hence will decouple from the low-energy theory. This procedure decouples (a subspace) of the $\frac{1}{2}$-BPS states of $\mathcal{N}=4$ super Yang–Mills theory. These low-lying states are protected by supersymmetry and will not be lifted from zero energy by interactions. In what follows, we assume that, even in the presence of interactions, these states remain decoupled (which amounts to assuming that interactions do not make any of the heavy states light). The complex scalar Z can be decomposed into partial waves on the S^3. The s-wave is simply a singlet under the $SO(4)$ symmetry of the S^3 on which the field theory is defined. The higher spherical harmonics have a greater energy and hence are among the states that decouple. We thus come to the remarkable conclusion that the limit we study is described exactly by the quantum mechanics of a single complex matrix. The action of $\mathcal{N}=4$ super Yang–Mills theory on $R\times S^3$ includes a mass term which arises from conformal coupling to the

[3] See [10] for closely related ideas.

metric of S^3. With a convenient normalization of the action, the free field theory propagators are

$$\langle Z^{\dagger}_{ij}(t) Z_{kl}(t)\rangle = \delta_{il}\delta_{jk} = \langle Y^{\dagger}_{ij}(t) Y_{kl}(t)\rangle = \langle X^{\dagger}_{ij}(t) X_{kl}(t)\rangle.$$

As long as one restricts attention to traces involving only Z or only $Z^{\dagger}$, it is possible to express the theory in terms of the eigenvalues of Z. The change of variables entailed in going from Z to the eigenvalues of Z induces a non-trivial Jacobian – the Van der Monde determinant. The net effect of this Jacobian is accounted for by treating the eigenvalues as fermions [11]. Consequently, one obtains the dynamics of N non-interacting fermions in an external harmonic oscillator potential.

Key idea: *The $\frac{1}{2}$-BPS sector of $\mathcal{N}=4$ super Yang–Mills theory is described exactly by the holomorphic sector of the quantum mechanics of a single complex matrix, which itself is equivalent to the dynamics of N free fermions.*

9.3 Dictionary

The $\frac{1}{2}$-BPS sector of Type IIB string theory on $\mathrm{AdS}_5\times\mathrm{S}^5$ contains gravitons, membranes and strings. Apparently all of these objects are captured by the matrix quantum mechanics of the previous section. To see that this is indeed plausible, recall that as the $\mathcal{R}$-charge (J) of an operator in the $\mathcal{N}=4$ super Yang–Mills theory is changed, its interpretation in the dual quantum gravity changes. This is a consequence of the Myers effect [12]: the background we are studying has a non-zero RR 5-form field strength switched on. This flux couples to D3 branes. Gravitons carry a D3 dipole charge and are hence polarized by the background flux [13]. As we increase J, the coupling to the background RR 5-form flux increases and the graviton expands. It puffs out to a radius

$$R = \sqrt{\frac{J}{N}}\, R_{\mathrm{AdS}}, \qquad \text{where} \qquad R^2_{\mathrm{AdS}} = \sqrt{g^2_{\mathrm{YM}} N}\,\alpha' .$$

We will consider the limit that N is very large with g^2_{YM} fixed and very small. Since the string coupling $g_s = g^2_{\mathrm{YM}}$, this is the weak string coupling and small curvature limit in which we expect to be able to recognize the familiar objects of perturbative string theory. For $J\sim O(1)$ the operator is dual to an object of zero size in string units, that is, a point-like graviton [1]. For $J\sim O(\sqrt{N})$ the operator is dual to an object of fixed size in string units – this is a string [10]. For $J\sim O(N)$ the operator is dual to an object whose size is of the order of R_{AdS} – as argued in [3, 14] these are the giant gravitons of [13]. Finally, consider $J\sim O(N^2)$. Naively, the size of these objects diverges, even when measured in units with $R_{\mathrm{AdS}}=1$. We will see that this divergence is simply an indication that these operators do not have an interpretation in terms of a new object in $\mathrm{AdS}_5\times\mathrm{S}^5$: these operators correspond to new backgrounds [5, 6].

The original $\mathcal{N}=4$ super Yang–Mills theory is defined on Minkowski space. After Wick rotating (to four-dimensional Euclidean space) and performing a conformal transformation, we obtain $\mathcal{N}=4$ super Yang–Mills theory on $R\times S^3$. Operators of the theory on four-dimensional Euclidean space are in one-to-one correspondence with states of the theory on $R\times S^3$, by the usual operator–state correspondence available for any conformal field theory. $R\times S^3$ is the boundary of $\mathrm{AdS}_5\times\mathrm{S}^5$ in global coordinates. It is natural to identify this boundary with the space on which the field theory lives. Taken together, we obtain a map between operators of the original $\mathcal{N}=4$ super Yang–Mills theory on Minkowski space and states of the string theory. For this reason we will often talk of "matching operators to states" and we will often be able to read the inner product between two states from a correlation function of the corresponding operators.

The symmetries of the $\mathcal{N}=4$ super Yang–Mills theory match the isometries present in the dual $\mathrm{AdS}_5\times\mathrm{S}^5$ background. When trying to match a specific operator to a specific state, it is useful to match the labels provided by these symmetries on both sides of the correspondence. Reasoning in this way, it is possible to argue that scaling dimensions in the field theory correspond to energies in the dual string theory and $\mathcal{R}$-charge in the field theory corresponds to angular momentum in the string theory.

Key idea: *The dictionary between the $\frac{1}{2}$-BPS sector of $\mathcal{N}=4$ super Yang–Mills theory and Type IIB string theory on $AdS_5\times S^5$ is organized according to $\mathcal{R}$-charge. When trying to match a specific operator to a specific state, it is useful to match scaling dimensions ($\mathcal{R}$-charge) in the field theory to energies (angular momentum) in the dual string theory.*

9.4 Organizing the degrees of freedom of a matrix model

In the previous section we have seen that, in order to capture all of the objects in the spectrum of the dual string theory, it is necessary to consider all possible values of the $\mathcal{R}$-charge. This is a complicated problem since the usual simplifications of the large-N limit are no longer present, as we now explain.

First consider the case that $n=O(1)$. A suitable basis is provided by single trace operators, $\mathrm{Tr}(Z^n)$. Normalize the operators with factors of N so that they have an $O(1)$ two-point function at large N:

$$\mathcal{O}_n=\frac{\mathrm{Tr}(Z^n)}{N^{\frac{n}{2}}}\,.$$

In this case we obviously have

$$\langle \mathcal{O}_n \mathcal{O}_m^\dagger \rangle \propto \delta_{mn} \,,$$

because the two operators have a different $\mathcal{R}$-charge if $m \neq n$. Consider next the correlator

$$\langle \mathcal{O}_p \mathcal{O}_n \mathcal{O}_{n+p}^\dagger \rangle \sim \frac{1}{N} \,.$$

The total $\mathcal{R}$-charge of this three-point function is zero, so it is not forced to vanish. However, recall that in the large-N limit the expectation values of observables factorize, which is equivalent to the statement that disconnected diagrams dominate. The leading (disconnected) contributions to the above correlator vanish and hence this correlator is suppressed in the large-N limit by the usual arguments. To explain why this correlator vanished, we can identify the number of traces with particle number – in which case the vanishing of the above correlator is the statement that although a two-particle state with gravitons of $\mathcal{R}$-charge p and n has the same $\mathcal{R}$-charge as a single-particle state with $\mathcal{R}$-charge $p+n$, the two states are orthogonal. The fact that the two states have a different particle number *explains* why their overlap is zero. Consequently, the weakly interacting supergravity Fock space is clearly visible in the dual gauge theory. There is a combinatoric coefficient on the right-hand side of the above correlator which counts the number of Wick contractions.

For $n = O(N)$, the usual $\frac{1}{N}$ suppression of non-planar diagrams is compensated by huge combinatoric factors,[4] so that operators composed of a product of a different number of traces are no longer orthogonal. Clearly then, the gravity states dual to single-trace operators are no longer orthogonal and hence there is no reason to expect that they will have a natural physical interpretation. The fact that these states are no longer orthogonal has a very natural explanation in the dual string theory. Recall that the dimension of the operator maps into the energy of the dual state. Thus, by considering operators of a very large dimension we are talking about very heavy objects in the dual string theory. As we increase the mass of the objects we study, we turn the gravitational interactions on and consequently the states that were orthogonal when there was no interaction are no longer orthogonal.

Ideally one needs a new basis in which the two-point functions are again diagonal. Corley, Jevicki and Ramgoolam have demonstrated that the Schur polynomials (in the zero coupling limit) do indeed diagonalize the two-point function [3] for the

[4] The number of Wick contractions explodes as more and more operators are included in each trace [14].

theory with gauge group $U(N)$. The Schur polynomial is defined by

$$\chi_R(Z) = \frac{1}{n!} \sum_{\sigma \in S_n} \chi_R(\sigma) \mathrm{Tr}(\sigma Z^{\otimes n}), \tag{9.1}$$

$$\mathrm{Tr}(\sigma Z^{\otimes n}) = Z^{i_1}_{i_{\sigma(1)}} Z^{i_2}_{i_{\sigma(2)}} \cdots Z^{i_{n-1}}_{i_{\sigma(n-1)}} Z^{i_n}_{i_{\sigma(n)}}.$$

The Schur polynomial label R can be thought of as a Young diagram which has n boxes. $\chi_R(\sigma)$ is the character of $\sigma \in S_n$ in representation R. For an extension of these results to the case of gauge group $SU(N)$, see [15].

Schur calculus

The dynamical content of any quantum theory is encoded in its correlation functions. Focusing on the $\frac{1}{2}$-BPS sector, Schur polynomials provide an excellent set of variables to probe different aspects of the dual string theory since (i) in the free field limit, the two-point functions of Schur polynomials are known exactly and (ii) they satisfy a nice product rule that can be used to collapse any *product* of Schur polynomials into a *sum* of polynomials. This product rule follows as a consequence of the Schur–Weyl duality between unitary groups and symmetric groups. As a consequence of the duality, Schur polynomials $\chi_R(U)$, when evaluated on an element $U \in U(N)$, give the character of U in the irreducible representation R. For any two such irreducible representations, R and S, it is well known that $R \otimes S = \oplus_T f_{RS;T} T$ where the $f_{RS;T}$ are known as Littlewood–Richardson numbers. With the interpretation of the Schur polynomials as characters, it follows immediately that

$$\chi_R(Z)\chi_S(Z) = \sum_T f_{RS;T} \chi_T(Z).$$

One immediate repercussion of this is that the *exact* computation of any multi-point extremal correlator of Schur polynomials can be collapsed down to an evaluation of two-point correlators [16].

There is a very natural connection between the free fermion description of Section 9.2 and the Schur polynomials [3,4]. The Schur polynomials are labeled by Young diagrams, which can be specified by giving a list of N integers, r_i, which count the number of boxes in the ith row of the Young diagram. The fermion wave function can be described by specifying the N occupied energy levels E_i, which is again a list of N integers. Detailed computations show that the Schur polynomials coincide with the N-fermion wave functions as long as we identify (see [3,4] for details)

$$E_i = N + i + r_i \, .$$

Thus, the Schur polynomial basis coincides with the free fermion basis.

Although the huge simplifications discussed above do not survive when one goes beyond the $\frac{1}{2}$-BPS sector, it is possible to write down more general bases which continue to diagonalize the two-point function [18–21]. These techniques were developed using crucial lessons [3,17] gained from the $\frac{1}{2}$-BPS sector. We will have more to say about these more general bases in the sections to come, since they are relevant for describing nearly supersymmetric states and hence they suffer only mild corrections.

Key idea: *The basis of the $\frac{1}{2}$-BPS sector of $\mathcal{N}=4$ super Yang–Mills theory provided by the Schur polynomials diagonalizes the free two-point function for any value of the $\mathcal{R}$-charge. At large $\mathcal{R}$-charge it will thus replace the trace basis, which now fails to diagonalize the free two-point function.*

9.5 Gravitons

In this section we will focus on that portion of the AdS/CFT dictionary that concerns operators with an $\mathcal{R}$-charge of $O(1)$. In this case, as explained above, the trace basis is perfectly acceptable and so we take $\mathcal{O}_n = \mathrm{Tr}(Z^n)$. We expect that these operators are dual to gravitons. In fact, this can be checked in detail, as we now explain.

The AdS/CFT correspondence claims that for every bulk field Φ in the gravitational description, there is a corresponding gauge-invariant operator O_ϕ. Asymptotically, AdS spaces have a boundary at spatial infinity and one needs to impose appropriate boundary conditions there. As a result, the partition function of the bulk theory is a functional of these boundary conditions. The boundary values of the fields are identified with sources that couple to the dual operator so that the gravitational partition function (the next formulas are schematic)

$$Z_{\text{gravity}}[\phi_0] \equiv \int_{\Phi|_{\partial(\text{AdS})}=\phi_0} D\Phi \; e^{-S}$$

is identified with the generating functional of correlation functions in the quantum field theory[5]

$$Z_{\text{gravity}}[\phi_0] = \left\langle e^{-\int \phi_0 O_\phi} \right\rangle_{\text{QFT}} .$$

5 The right-hand side of this relation suffers from the usual UV divergences present in any quantum field theory, and hence needs to be renormalized. The left-hand side suffers from IR divergences and hence also requires renormalization. The details of this renormalization have been worked out in [22,23].

Since the gravitons are meant to be dual to operators with $\mathcal{R}$-charge of $O(1)$, and since graviton dynamics is captured by the supergravity approximation to the complete string theory, using the above relation we should be able to compute the correlation functions of the O_n at strong coupling and at large N. Further, since these operators enjoy some protection against corrections by virtue of their supersymmetry, we may be optimistic that the strong-coupling and weak-coupling results will agree. The computation has been performed and the agreement is perfect [24].

Key idea: *The identification of gravitons with operators of $\mathcal{R}$-charge of O(1) can be checked by using the AdS/CFT correspondence. The agreement is perfect.*

9.6 Strings

We now move on to operators with an $\mathcal{R}$-charge of $J = O(\sqrt{N})$. These objects are already heavy enough that, for single traces, we have a new effective string coupling replacing $\frac{1}{N}$:

$$g_s \sim \frac{1}{N} \to \frac{J^2}{N}\,.$$

Thus, to suppress non-planar corrections we need to take $J^2 \ll N$, which we assume from now on. To see stringy physics it is useful to consider operators, for example, of the form

$$\mathrm{Tr}(YZ^nYZ^{J-n})\,.$$

Due to the presence of the Y fields, this state has $\Delta = J+2$ and $\mathcal{R}$-charge J. Since $J^2 \to \infty$, this operator is nearly $\frac{1}{2}$-BPS and we might still expect that corrections are suppressed. This is indeed the case [10]: one finds that the expansion parameter λ is replaced by $\frac{\lambda}{J^2}$. The eigenvalues of the dilatation operator when acting on this class of states precisely matches the expected energies of strings in the dual string theory [10]. The matching of spectra can be significantly improved. Think of the Ys and Zs as populating a lattice with $J+2$ sites. Further, identify the Ys with a spin-up state and the Zs with a spin-down state. With this interpretation, the Yang–Mills dilatation operator can be identified with the Hamiltonian of a spin chain [25–28]. By considering coherent states of the spin chain, one can give the spin chain a sigma model description. The resulting sigma model agrees exactly with the sigma model for a string rotating with a large angular momentum, so that now agreement is obtained at the level of the action [29]. The very detailed agreement allows one to frame very precise questions. For example, it is straightforward to show that the mean value of the spin of the spin chain corresponds to the position of the string in the dual gravitational spacetime (see [29] for details).

Key idea: *Operators with an $\mathcal{R}$-charge of $O(\sqrt{N})$ do indeed correspond to strings. The stringy dynamics can be recovered from the field theory and further, it is clear* *how to build field theory states corresponding to strings localized at a point in the dual (higher-dimensional!) gravitational theory.*

9.7 Giant gravitons

In this section we study certain membranes in string theory, known as giant gravitons. These operators have an $\mathcal{R}$-charge of $O(N)$; in this case the trace basis badly fails to provide orthogonal states and hence we have the first case in which the Schur polynomials *must* be used.

Giant graviton solutions describe branes extended in the sphere [13] or in the AdS space [30, 31] of the AdS$\times$S background. The giant gravitons are (classically) stable[6] due to the presence of the five form flux which produces a force that exactly balances their tension. The force which balances the tension is a Lorentz-like[7] force so that the force increases with increased giant angular momentum. Consequently, a giant graviton expands to a radius proportional to the square root of its angular momentum [13]. If the giant is expanding in the S^5 of the AdS$\times$S background, there is a limit on how large it can be – its radius must be less than the radius of the S^5 [13]. This in turn implies a cut-off on the angular momentum of the giant. Since angular momentum of the giant maps into $\mathcal{R}$-charge, there should be a cut-off on the $\mathcal{R}$-charge of the dual operators. The Schur polynomials corresponding to totally antisymmetric representations do have a cut-off on their $\mathcal{R}$-charge; this cut-off exactly matches the cut-off on the giant's angular momentum [14]. Thus, it is natural to identify Schur polynomials for the completely antisymmetric representations as the operator dual to sphere giant gravitons. Another class of Schur polynomials which are naturally singled out are those corresponding to totally symmetric representations. Since these representations are not cut off, they are naturally identified as operators dual to AdS giant gravitons [3], which, because they expand in the AdS space, can expand to an arbitrarily large size and hence have no bound on their angular momentum. The Young diagram has at most N rows, implying a cut-off on the number of AdS giant gravitons; the need for this cut-off is again visible in the dual gravitational theory: it ensures that the 5-form flux at the center of the AdS space does not become zero; a non-zero flux is needed to support an AdS giant.

A smoking gun signal for D-branes in string theory is the presence of open strings in the excitation spectrum of the D-brane. The low-energy dynamics of these strings should realize a new gauge theory on the brane worldvolume. Since the brane worldvolume is a compact space, Gauss' Law will only allow excitations with vanishing

[6] They carry an RR-dipole charge, but no RR-monopole charge. Thus their decay would not be prevented by charge conservation.
[7] That is, velocity-dependent.

total charge. There is a very natural generalization of the Schur polynomials, the restricted Schur polynomials, in which the symmetric group character appearing in (9.1) is replaced by a partial trace over the symmetric group elements. The number of partial traces that can be defined matches the number of states consistent with the Gauss Law constraint [8]. Surprisingly, it is still possible to compute correlators of restricted Schur polynomials and to evaluate the action of the field theory dilatation operator on these operators [32–34]. This is rather non-trivial: due to the large $\mathcal{R}$-charge of the operators one needs to sum an infinite number of non-planar diagrams in these computations. The dilatation operator can again be matched to the action for open strings, at the level of the action [33]. This allows one to ask questions that could not be asked in perturbative string theory. For example, when the string and membrane interact, the Young diagram changes shape. This allows one to take backreaction on the membrane into account [33, 34]. From the point of view of perturbative string theory, the D-brane appears as a boundary condition and it is not obvious how one should account for backreaction. Another interesting effect discovered by studying these operators is an instability arising when long open strings are attached to the giant graviton [33]. The giant graviton, which couples to the RR 5-form flux, does not undergo geodesic motion. The open string, which does not couple to the RR 5-form flux, would like to undergo geodesic motion but is being dragged in a non-geodesic motion by the giant. These centrifugal forces can overcome the string tension if the string is long enough, leading to the instability. This effect has also been exploited as a toy model for quantum gravity effects in braneworld cosmological models [35].

Emergent gauge theory

Fundamental strings are charged under the Kalb–Ramond 2-form. For stretched strings this charge can be thought of as a current flowing along the string. Consequently, when they end on D-branes, conservation of string charge means that strings act as “sources” or “sinks” on the brane worldvolume. When – as in the case of the giant gravitons above – the D-brane wraps a compact space like the $S^3 \subset S^5$, the only allowed attached open-string states are those consistent with Gauss’ Law. So, for example, for the giant configurations in Figure 9.1 only (b) is a valid state. This is a manifestation of the closed topology of the spherical D-brane. To see how Gauss’ Law is encoded in the Yang–Mills theory, we need to consider operators dual to *excited* giant gravitons. These *restricted* Schur polynomials,

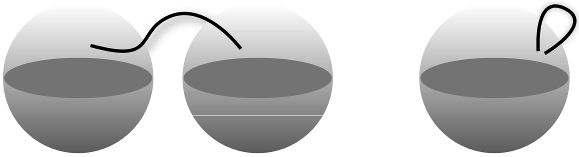


Figure 9.1 Giants with strings attached

$$\chi^{(k)}_{R,R_1} = \frac{1}{(n-k)!}\sum_{\sigma\in S_n} \mathrm{Tr}_{R_1}\left(\Gamma_R(\sigma)\right)\times$$
$$\mathrm{Tr}\left(\sigma Z^{\otimes n-k}\left(W^{(1)}\right)^{i_{n-k+1}}_{i_{\sigma(n-k+1)}}\cdots\left(W^{(k)}\right)^{i_n}_{i_{\sigma(n)}}\right),$$

are obtained from (9.1) by the insertion of "words" $W^{(i)}$ describing the open strings attached to the giant system. In the language of Young diagrams, this corresponds to labeling boxes in the diagram associated to the giant graviton. For the single string attached to a single sphere giant graviton with momentum p in Figure 9.1(b), for example, the restricted Schur takes the form $\chi^{(1)}_{\mathbf{1}^{p+1},\mathbf{1}^p}(Z,W)$ with the open string word $W^i_j = (Y^J)^i_j$ say. We have denoted a Young diagram with a single column containing n boxes by 1^n. To match with the gravity side of the correspondence, we need to count the number of possible operators $\chi^{(k)}_{R,R_1}(Z,W^i)$ that can be constructed for a given representation. Remarkably, the number of these operators matches precisely with the number of allowed states in the string theory that satisfy the Gauss constraint [8].

Key idea: *Giant gravitons are membrane states that are represented in the dual field theory by Schur polynomials and restricted Schur polynomials. The field theory correctly reproduces both the number of open-string excitations of the giant gravitons and their dynamics.*

9.8 New geometries

When the operator we consider has an $\mathcal{R}$-charge of $O(N^2)$, we are studying objects in the dual string theory that are so heavy that they backreact on spacetime. Consequently, there should be new supergravity backgrounds corresponding to these operators. These new backgrounds should preserve an $R \times SO(4) \times SO(4)$ symmetry. The conserved charge associated to R corresponds to conformal dimension in the field theory. The first $SO(4)$ symmetry corresponds to rotations of the S^3 on which the field theory is defined (recall that our operators are built from the s-wave of Z). The second $SO(4)$ corresponds to $\mathcal{R}$-symmetry rotations of the four scalars $\phi_3, \phi_4, \phi_5, \phi_6$ which are not used to construct Z. This ansatz is sufficiently specific that the general solution with these isometries can be written down. These geometries, constructed by Lin, Lunin and Maldacena, are known as the LLM geometries [5]. The complete solution is determined in terms of a single function which obeys a three-dimensional Laplace equation. To get a unique solution, one needs to specify a boundary condition for this function on a specific two-dimensional plane. In order that the geometry is regular, this boundary condition must assign the function either of the values $\pm\frac{1}{2}$ on this plane. It is tempting to identify this two-dimensional plane as the phase space of the fermions of Section 9.2, where regions with $\frac{1}{2}$ correspond to occupied states and regions with $-\frac{1}{2}$ correspond to unoccupied states [5]. This can indeed be checked in detailed computations and it turns out that this identification is perfect [36]. Thus, once again the quantum field theory and the quantum gravity theories are in complete agreement.

Denote the Schur polynomial of $\mathcal{R}$-charge of $O(N^2)$ by $\chi_B(Z)$. The operator $\chi_B(Z)$ creates the new background. Is it possible, directly in the field theory, to construct the metric corresponding to this dual geometry? One thing we could do is follow the propagation of a graviton in the spacetime. Since we know how to build an operator dual to a graviton, this is a computation we know how to do. Further, gravitons move along null geodesics and hence they "know" about the dual geometry, so that we might indeed hope to learn something about the metric. To create a graviton moving on the new background we need to act with $\mathrm{Tr}(Z^n)$ with n a number of $O(1)$. Thus, to probe the geometry we need to compute the correlator

$$\left\langle \chi_B(Z)\chi_B(Z)^\dagger \mathrm{Tr}(Z^n)\mathrm{Tr}(Z^n)^\dagger \right\rangle .$$

This computation can be performed exactly, in the free field limit, for any choice of the Young diagram B [37, 38]. As soon as B has $O(N^2)$ boxes, one can no longer neglect Wick contractions between fields in $\mathrm{Tr}(Z^n)^\dagger$ and fields in the background $\chi_B(Z)$ – at precisely the value of the $\mathcal{R}$-charge that we expect backreaction is important, the graviton and background start interacting! This is how the field theory accounts for back reaction [37,38]. The resulting correlators are surprisingly simple. If one takes B to be a Young diagram with M columns (M is $O(N)$), the net effect of the background is to replace $N \to M + N$. Recall that in the trivial background graviton correlators admit a $\frac{1}{N}$ expansion. In the new background, graviton correlators are organized by a $\frac{1}{M+N}$ expansion [37, 38]. This renormalization of the string coupling constant was achieved by summing an infinite number of non-planar diagrams, something that is only possible thanks to the power of the Schur polynomial technology. This renormalization of the string coupling can be checked rather explicitly, using holography in the LLM background [39], along the lines of the computation of Section 9.5. One again finds perfect agreement [40]. Apart from probing the geometry with gravitons, one could consider probing the geometry with strings [41, 42] or even giant gravitons [43]. This leads to some interesting results. For example, in the case of string probes, one can again construct the dilatation operator and study the sigma model that arises from the coherent state expectation value of the model. In this way it is possible to read off the metric that the string feels[8] [41, 42], which is rather detailed information about the dual geometry.

Given this very concrete description of the $\frac{1}{2}$-BPS geometries, it is possible to re-examine some long-standing puzzles. An important problem in this class is the information loss paradox. The entropy of black holes suggests an enormous degeneracy of microscopic states. The information loss paradox would be evaded if one could show that a pure initial state collapses to a particular pure black hole microstate whose exact structure can be deduced by careful measurements. What do pure microstates look like and what sorts of measurements can distinguish them from each other? In [6], this problem was examined by applying information theoretic ideas to Schur polynomials with $O(N^2)$ boxes. It is possible to characterize a "typical operator" and then to ask what the semiclassical description, in the dual gravity, of this state is. A typical very heavy state corresponds to a spacetime "foam." Almost no semiclassical probes will be able to distinguish different foam states, and the resulting effective description gives a singular geometry. Although the $\frac{1}{2}$-BPS states considered are not black holes, this study seems to explain how the existence of pure underlying microstates and the absence of fundamental information loss are consistent with the thermodynamic character of semiclassical black holes [6].

Local gravitons

For nearly a century now, we have learnt that to understand the geometry of spacetime, we have to probe it with localized objects; (very roughly) if you want to know whether a spacetime is curved or not, throw some particles into it and chart their geodesics. From a field-theorist's point of view, local geometric structure in spacetime arises as a coherent excitation of gravitons. To understand how bulk spacetime

[8] The string only moves close to the plane on which the boundary condition is specified. One is able to read off the metric on this plane, which is what we mean by "the metric that the string feels." Of course, one would like to do better and determine the full metric.

geometry *emerges* in the Yang–Mills matrix model, one might try to understand the localization and dynamics of gravitons.

To this end, following the arguments outlined above, the normalized operator dual to a graviton of one unit of angular momentum (in S^5) is $\frac{\mathrm{Tr}(Z)}{\sqrt{N}}$. Identifying this operator with the graviton creation operator $a^\dagger$ in the dual quantum gravity theory, we can also define a graviton annihilation operator $\frac{1}{\sqrt{N}}\mathrm{Tr}\left(\frac{d}{dZ}\right) \leftrightarrow a$, so that $\left[a, a^\dagger\right] = 1$. From this, a graviton coherent state operator can be built in the usual way,

$$\mathcal{O}_z = e^{-\frac{1}{2}|z|^2} e^{z\frac{\mathrm{Tr}(Z)}{\sqrt{N}}} .$$

After conformal mapping to $R \times S^3$, the operator $\mathcal{O}_z$ is mapped into the coherent state $|z\rangle$ with $z = r\, e^{-i\phi}$. The dynamics of the low-energy excitations of this coherent state on $R \times S^3$ is captured by the Landau–Lifshitz Lagrangian

$$L = \langle z|i\frac{d}{dt}|z\rangle - \langle z|H|z\rangle = \dot{\phi} r^2 - r^2,$$

when evaluated on the graviton coherent state. The equations of motion can be integrated to determine $\dot{\phi} = 1$ and $\dot{r} = 0$, which are by now familiar results for gravitons in $\mathrm{AdS}_5 \times \mathrm{S}^5$. These results are not unique to $\mathrm{AdS}_5 \times \mathrm{S}^5$ and can be extended to a whole class of $\frac{1}{2}$-BPS geometries – the so-called LLM backgrounds which are built from operators with $\mathcal{R}$-charge of $\mathcal{O}(N^2)$. These results can be extended to describe gravitons localized at different values of r in the multi-ring LLM geometries providing compelling evidence that *local* geometry emerges from the super Yang–Mills gauge theory [37].

Key idea: *The field theory description of $\frac{1}{2}$-BPS geometries is in terms of operators with dimension of $O(N^2)$. Probing this geometry corresponds to computing correlators of operators with the background operator inserted.*

9.9 Outlook

It is clear that something non-trivial is working. Things are, however, far from satisfactory. The AdS/CFT correspondence has passed highly non-trivial tests (far more than we discussed) so that we are confident the basic idea is correct. However, we still have no real understanding of why the degrees of freedom of a strongly coupled Yang–Mills theory are most simply described starting from a higher-dimensional dynamical geometry. The geometry is not visible in the weak coupling (Lagrangian) description of the quantum field theory, and in this sense is emergent. A simple example of an emergent geometry is provided by the large-N limit of a single Hermitian matrix model quantum mechanics [2]. In the large-N limit, the integral is dominated by a saddle point with a definite eigenvalue distribution. The emergent geometry relies crucially on the repulsive inter-eigenvalue force which causes the eigenvalues to spread. Without this repulsive force the eigenvalues would simply sit at one of the minima of the potential of the matrix model. The repulsive force itself comes from the Van der Monde determinant, that is, from the measure of the path integral.[9] It is thus a quantum effect, which seems to match nicely with the fact that $\hbar$ sets the radius of the AdS space (see footnote 1). To go beyond a single matrix is extremely difficult. For a single-matrix model, the eigenvalues provide a very convenient set of variables. For more than one matrix, since the matrices will not in general commute, it's not clear what the analog of the eigenvalues are. An important fact might be that we only expect the emergence of geometry in the strong coupling limit. Berenstein [7, 45] has suggested that in the strong-coupling and large-N limit the matrices will commute with each other. In this way the usual N^2 degrees of freedom of matrices get effectively reduced to order N degrees of freedom and the collective description of these low-energy degrees of freedom can be given in terms of a joint eigenvalue distribution for several matrices. It is the geometrical description of this eigenvalue distribution that is supposed to produce the emergent geometry [7, 45].

The emergent geometry obtained in the large-N limit of a single Hermitian matrix model quantum mechanics [2] was constructed by making a systematic change of variables in the quantum theory, correctly accounting for the Jacobian of this change. A systematic way to achieve this is provided by the collective field theory formalism [46]. For a single matrix a convenient set of gauge-invariant variables is provided by traces of powers of the matrix. For more than one matrix one needs to consider traces of arbitrary words built from the matrices – which is rather complicated. However, such a rewriting does indeed show that the field theory reconstructs the dual gravitational dynamics [47]. In principle this approach would provide a way to systematically reconstruct the dual gravitational dynamics from the field theory. The problem, however, is one of considerable complexity. It seems that additional insight is needed before this program can be taken to completion.

[9] The Van der Monde determinant has recently been shown to arise in certain sectors of multi-matrix models [44].

Above we have restricted ourselves to the $\frac{1}{2}$-BPS sector. A skeptic might suggest that it is dangerous to draw general lessons from the $\frac{1}{2}$-BPS sector. Indeed it might be, so we should try to do better. To go beyond the $\frac{1}{2}$-BPS sector, one needs to study multi-matrix dynamics. In general, this is a formidable problem. There has, however, been some recent progress: three independent bases for general multi-matrix models have been identified. For a review of these developments and the work leading up to them, see [9]. The basis described in [18] builds operators with definite flavor quantum numbers. The basis of [19] uses the Brauer algebra to build correlators involving Z and $Z^{\dagger}$; this basis seems to be the most natural for exploring brane/antibrane systems. The basis of [20] most directly allows one to consider excitations of the operator. All three bases diagonalize the two-point functions in the free field theory limit (to all orders in the $\frac{1}{N}$ expansion); it is in this sense that they generalize the Schur polynomials to the case of multi-matrix models. The discovery of these bases seems to be a promising start towards exploring $\mathcal{N}=4$ super Yang–Mills theory beyond the $\frac{1}{2}$-BPS sector. On the gravity side there has been some progress in characterizing the $\frac{1}{4}$-BPS geometries [48]. Perhaps in the not too distant future we will have an answer for George.

Acknowledgments

We would like to thank Rajsekhar Bhattacharyya, Tom Brown, Storm Collins, Tanay Dey, George Ellis, Alex Hamilton, Norman Ives, Antal Jevicki, Yusuke Kimura, Andrea Prinsloo, Sanjaye Ramgoolam, Joao Rodrigues, Jelena Smolic, Milena Smolic, Michael Stephanou, Dave Turton, Nick Toumbas and Alex Welte for pleasant discussions and/or helpful correspondence. This work is based upon research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation. JM is supported by the National Research Foundation under the Thuthuka and Key International Scientific Collaboration programs. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors and therefore the NRF and DST do not accept any liability with regard thereto.

References

[1] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," *Adv. Theor. Math. Phys.* **2**, 231 (1998) [*Int. J. Theor. Phys.* **38**, 1113 (1999)] [arXiv:hep-th/9711200]; S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from non-critical string theory," *Phys. Lett. B* **428**, 105 (1998) [arXiv:hep-th/9802109]; E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2**, 253 (1998) [arXiv:hep-th/9802150].

[2] S. R. Das and A. Jevicki, "String field theory and physical interpretation of $D = 1$ Strings," *Mod. Phys. Lett. A* **5**, 1639 (1990).
[3] S. Corley, A. Jevicki, and S. Ramgoolam, "Exact correlators of giant gravitons from dual $N = 4$ SYM theory," *Adv. Theor. Math. Phys.* **5**, 809 (2002) [arXiv:hep-th/0111222].
[4] D. Berenstein, "A toy model for the AdS/CFT correspondence," *JHEP* **0407**, 018 (2004) [arXiv:hep-th/0403110].
[5] H. Lin, O. Lunin, and J. M. Maldacena, "Bubbling AdS space and 1/2 BPS geometries," *JHEP* **0410**, 025 (2004) [arXiv:hep-th/0409174].
[6] V. Balasubramanian, J. de Boer, V. Jejjala, and J. Simon, "The library of Babel: On the origin of gravitational thermodynamics," *JHEP* **0512**, 006 (2005) [arXiv:hep-th/0508023]; V. Balasubramanian, V. Jejjala, and J. Simon, "The library of Babel," *Int. J. Mod. Phys. D* **14**, 2181 (2005) [arXiv:hep-th/0505123].
[7] D. Berenstein, "Large N BPS states and emergent quantum gravity," *JHEP* **0601**, 125 (2006) [arXiv:hep-th/0507203].
[8] V. Balasubramanian, D. Berenstein, B. Feng, and M. x. Huang, "D-branes in Yang–Mills theory and emergent gauge symmetry," *JHEP* **0503**, 006 (2005) [arXiv:hep-th/0411205].
[9] S. Ramgoolam, "Schur–Weyl duality as an instrument of gauge–string duality," arXiv:0804.2764 [hep-th].
[10] D. Berenstein, J. M. Maldacena, and H. Nastase, "Strings in flat space and pp waves from $N = 4$ super Yang–Mills," *JHEP* **0204**, 013 (2002) [arXiv:hep-th/0202021].
[11] E. Brezin, C. Itzykson, G. Parisi, and J. B. Zuber, "Planar diagrams," *Commun. Math. Phys.* **59**, 35 (1978).
[12] R. C. Myers, "Dielectric-branes," *JHEP* **9912**, 022 (1999) [arXiv:hep-th/9910053].
[13] J. McGreevy, L. Susskind, and N. Toumbas, "Invasion of the giant gravitons from anti-de Sitter space," *JHEP* **0006**, 008 (2000) [arXiv:hep-th/0003075].
[14] V. Balasubramanian, M. Berkooz, A. Naqvi, and M. J. Strassler, "Giant gravitons in conformal field theory," *JHEP* **0204**, 034 (2002) [arXiv:hep-th/0107119].
[15] R. de Mello Koch and R. Gwyn, "Giant graviton correlators from dual $SU(N)$ super Yang–Mills theory," *JHEP* **0411**, 081 (2004) [arXiv:hep-th/0410236]; T. W. Brown, "Half-BPS $SU(N)$ correlators in $N = 4$ SYM," arXiv:hep-th/0703202.
[16] T. W. Brown, R. de Mello Koch, S. Ramgoolam, and N. Toumbas, "Correlators, probabilities and topologies in $N = 4$ SYM," *JHEP* **0703**, 072 (2007) [arXiv:hep-th/0611290].
[17] S. Corley and S. Ramgoolam, "Finite factorization equations and sum rules for BPS correlators in $N = 4$ SYM theory," *Nucl. Phys. B* **641**, 131 (2002) [arXiv:hep-th/0205221].
[18] T. W. Brown, P. J. Heslop, and S. Ramgoolam, "Diagonal multi-matrix correlators and BPS operators in $N = 4$ SYM," arXiv:0711.0176 [hep-th]; T. W. Brown, "Permutations and the loop," *JHEP* **0806**, 008 (2008) [arXiv:0801.2094 [hep-th]]; T. W. Brown, P. J. Heslop, and S. Ramgoolam, "Diagonal free field matrix correlators, global symmetries and giant gravitons," *JHEP* **0904**, 089 (2009) [arXiv:0806.1911 [hep-th]].
[19] Y. Kimura and S. Ramgoolam, "Branes, anti-branes and Brauer algebras in gauge–gravity duality," arXiv:0709.2158 [hep-th]; Y. Kimura, "Non-holomorphic multi-matrix gauge invariant operators based on Brauer algebra," arXiv:0910.2170 [hep-th].
[20] R. Bhattacharyya, S. Collins, and R. de Mello Koch, "Exact multi-matrix correlators," arXiv:0801.2061 [hep-th]; R. Bhattacharyya, R. de Mello Koch, and M. Stephanou, "Exact multi-restricted Schur polynomial correlators," *JHEP* **0806**, 101 (2008)

[arXiv:0805.3025 [hep-th]]; S. Collins, "Restricted Schur polynomials and finite N counting," *Phys. Rev. D* **79**, 026002 (2009) [arXiv:0810.4217 [hep-th]].

[21] Y. Kimura and S. Ramgoolam, "Enhanced symmetries of gauge theory and resolving the spectrum of local operators," *Phys. Rev. D* **78**, 126003 (2008) [arXiv:0807.3696 [hep-th]].

[22] M. Bianchi, D. Z. Freedman, and K. Skenderis, "How to go with an RG flow," *JHEP* **0108**, 041 (2001) [arXiv:hep-th/0105276]; M. Bianchi, D. Z. Freedman, and K. Skenderis, "Holographic renormalization," *Nucl. Phys. B* **631**, 159 (2002) [arXiv:hep-th/0112119]; K. Skenderis, "Lecture notes on holographic renormalization," *Class. Quant. Grav.* **19**, 5849 (2002) [arXiv:hep-th/0209067].

[23] K. Skenderis and M. Taylor, "Kaluza–Klein holography," *JHEP* **0605**, 057 (2006) [arXiv:hep-th/0603016]; K. Skenderis and M. Taylor, "Anatomy of bubbling solutions," *JHEP* **0709**, 019 (2007) [arXiv:0706.0216 [hep-th]].

[24] S. Lee, S. Minwalla, M. Rangamani, and N. Seiberg, "Three-point functions of chiral operators in $D=4$, $N=4$ SYM at large N," *Adv. Theor. Math. Phys.* **2**, 697 (1998) [arXiv:hep-th/9806074]; K. A. Intriligator, "Bonus symmetries of N = 4 super-Yang-Mills correlation functions via AdS duality," *Nucl. Phys. B* **551**, 575 (1999) [arXiv:hep-th/9811047]; B. U. Eden, P. S. Howe, A. Pickering, E. Sokatchev and P. C. West, "Four-point functions in N = 2 superconformal field theories," *Nucl. Phys. B* **581**, 523 (2000) [arXiv:hep-th/0001138]; B. U. Eden, P. S. Howe, E. Sokatchev and P. C. West, "Extremal and next-to-extremal n-point correlators in four-dimensional SCFT," *Phys. Lett. B* **494**, 141 (2000) [arXiv:hep-th/0004102].

[25] C. Kristjansen, J. Plefka, G. W. Semenoff and M. Staudacher, "A new double-scaling limit of $N=4$ super Yang–Mills theory and PP-wave strings," *Nucl. Phys. B* **643**, 3 (2002) [arXiv:hep-th/0205033]; N. R. Constable, D. Z. Freedman, M. Headrick, S. Minwalla, L. Motl, A. Postnikov and W. Skiba, "PP-wave string interactions from perturbative Yang–Mills theory," *JHEP* **0207**, 017 (2002) [arXiv:hep-th/0205089].

[26] J. A. Minahan and K. Zarembo, "The Bethe-ansatz for $N=4$ super Yang–Mills," *JHEP* **0303**, 013 (2003) [arXiv:hep-th/0212208].

[27] N. Beisert and M. Staudacher, "The $N=4$ SYM Integrable Super Spin Chain," *Nucl. Phys. B* **670**, 439 (2003) [arXiv:hep-th/0307042].

[28] N. Beisert, C. Kristjansen and M. Staudacher, "The dilatation operator of $N=4$ super Yang–Mills theory," *Nucl. Phys. B* **664**, 131 (2003) [arXiv:hep-th/0303060].

[29] M. Kruczenski, "Spin chains and string theory," *Phys. Rev. Lett.* **93**, 161602 (2004) [arXiv:hep-th/0311203]. M. Kruczenski, A. V. Ryzhov and A. A. Tseytlin, "Large spin limit of AdS(5) × S**5 string theory and low energy expansion of ferromagnetic spin chains," *Nucl. Phys. B* **692**, 3 (2004) [arXiv:hep-th/0403120].

[30] M. T. Grisaru, R. C. Myers and O. Tafjord, "SUSY and Goliath," *JHEP* **0008**, 040 (2000) [arXiv:hep-th/0008015].

[31] A. Hashimoto, S. Hirano and N. Itzhaki, "Large branes in AdS and their field theory dual," *JHEP* **0008**, 051 (2000) [arXiv:hep-th/0008016].

[32] R. de Mello Koch, J. Smolic and M. Smolic, "Giant Gravitons - with Strings Attached (I)," *JHEP* **0706**, 074 (2007), arXiv:hep-th/0701066.

[33] D. Berenstein, D. H. Correa and S. E. Vazquez, "A study of open strings ending on giant gravitons, spin chains and integrability," [arXiv:hep-th/0604123]; D. Berenstein, D. H. Correa and S. E. Vazquez, "Quantizing open spin chains with variable length: An example from giant gravitons," *Phys. Rev. Lett.* **95**, 191601 (2005) [arXiv:hep-th/0502172]; D. H. Correa and G. A. Silva, "Dilatation operator and the super Yang–Mills duals of open strings on AdS giant gravitons," *JHEP* **0611**, 059 (2006) [arXiv:hep-th/0608128].

[34] R. de Mello Koch, J. Smolic and M. Smolic, "Giant gravitons – with strings attached (II)," *JHEP* **0709** 049 (2007) [arXiv:hep-th/0701067]; D. Bekker, R. de Mello Koch and M. Stephanou, "Giant gravitons – with strings attached (III)," *JHEP* **0802**, 029 (2008) [arXiv:0710.5372 [hep-th]].

[35] A. Hamilton and J. Murugan, "On the shoulders of giants - quantum gravity and braneworld stability," [arXiv:0806.3273 [gr-qc]]

[36] L. Grant, L. Maoz, J. Marsano, K. Papadodimas and V. S. Rychkov, "Minisuperspace quantization of 'bubbling AdS' and free fermion droplets," *JHEP* **0508**, 025 (2005) [arXiv:hep-th/0505079]; L. Maoz and V. S. Rychkov, "Geometry quantization from supergravity: The case of 'bubbling AdS'," *JHEP* **0508**, 096 (2005) [arXiv:hep-th/0508059].

[37] R. de Mello Koch, "Geometries from Young diagrams," *JHEP* **0811**, 061 (2008) [arXiv:0806.0685 [hep-th]].

[38] R. de Mello Koch, N. Ives, and M. Stephanou, "Correlators in nontrivial backgrounds," *Phys. Rev. D* **79**, 026004 (2009) [arXiv:0810.4041 [hep-th]].

[39] K. Skenderis and M. Taylor, "Anatomy of bubbling solutions," *JHEP* **0709**, 019 (2007) [arXiv:0706.0216 [hep-th]].

[40] R. de Mello Koch, T. K. Dey, N. Ives, and M. Stephanou, "Correlators of operators with a large R-charge," arXiv:0905.2273 [hep-th].

[41] S. E. Vazquez, "Reconstructing 1/2 BPS space-time metrics from matrix models and spin chains," *Phys. Rev. D* **75**, 125012 (2007) [arXiv:hep-th/0612014].

[42] H. Y. Chen, D. H. Correa, and G. A. Silva, "Geometry and topology of bubble solutions from gauge theory," *Phys. Rev. D* **76**, 026003 (2007) [arXiv:hep-th/0703068].

[43] G. Mandal, "Fermions from half-BPS supergravity," *JHEP* **0508**, 052 (2005) [arXiv:hep-th/0502104].

[44] M. Masuku and J. P. Rodrigues, "Laplacians in polar matrix coordinates and radial fermionization in higher dimensions," arXiv:0911.2846 [hep-th]; Y. Kimura, S. Ramgoolam, and D. Turton, "Free particles from Brauer algebras in complex matrix models," arXiv:0911.4408 [hep-th].

[45] D. Berenstein, "A strong coupling expansion for $N = 4$ SYM theory and other SCFT's," arXiv:0804.0383 [hep-th]; D. E. Berenstein and S. A. Hartnoll, "Strings on conifolds from strong coupling dynamics: quantitative results," *JHEP* **0803** (2008) 072 [arXiv:0711.3026 [hep-th]]; D. Berenstein, "Strings on conifolds from strong coupling dynamics, part I," *JHEP* **0804** (2008) 002 [arXiv:0710.2086 [hep-th]]; D. E. Berenstein, M. Hanada, and S. A. Hartnoll, "Multi-matrix models and emergent geometry," *JHEP* **0902**, 010 (2009) [arXiv:0805.4658 [hep-th]].

[46] A. Jevicki and B. Sakita, "The quantum collective field method and its application to the planar limit," *Nucl. Phys. B* **165**, 511 (1980); A. Jevicki and B. Sakita, "Collective field approach to the large N limit: Euclidean field theories," *Nucl. Phys. B* **185**, 89 (1981).

[47] J. P. Rodrigues, "Large N spectrum of two matrices in a harmonic potential and BMN energies," *JHEP* **0512**, 043 (2005) [arXiv:hep-th/0510244]; A. Donos, A. Jevicki, and J. P. Rodrigues, "Matrix model maps in AdS/CFT," *Phys. Rev. D* **72**, 125009 (2005) [arXiv:hep-th/0507124]; R. de Mello Koch, A. Jevicki, and J. P. Rodrigues, "Instantons in $c = 0$ CSFT," *JHEP* **0504**, 011 (2005) [arXiv:hep-th/0412319]; R. de Mello Koch, A. Donos, A. Jevicki, and J. P. Rodrigues, "Derivation of string field theory from the large N BMN limit," *Phys. Rev. D* **68**, 065012 (2003) [arXiv:hep-th/0305042]; R. de Mello Koch, A. Jevicki, and J. P. Rodrigues, "Collective string field theory of matrix models in the BMN limit," *Int. J. Mod. Phys. A* **19**, 1747 (2004) [arXiv: hep-th/0209155].

[48] A. Donos, "A description of 1/4 BPS configurations in minimal type IIB SUGRA," *Phys. Rev. D* **75**, 025010 (2007) [arXiv:hep-th/0606199]; B. Chen *et al.*, "Bubbling AdS and droplet descriptions of BPS geometries in IIB supergravity," *JHEP* **0710**, 003 (2007) [arXiv:0704.2233 [hep-th]]; O. Lunin, "Brane webs and 1/4-BPS geometries," arXiv:0802.0735 [hep-th].

10

Loop quantum gravity

HANNO SAHLMANN

In this chapter we review the foundations and present status of loop quantum gravity. It is short and relatively non-technical, the emphasis is on the ideas, and the flavor of the techniques. In particular, we describe the kinematical quantization and the implementation of the Hamilton constraint, as well as the quantum theory of black hole horizons, semiclassical states, and matter propagation. Spin foam models and loop quantum cosmology are mentioned only in passing, as these will be covered in separate reviews to be published alongside this one.

10.1 Introduction

Loop quantum gravity is a non-perturbative approach to the quantum theory of gravity, in which no classical background metric is used. In particular, its starting point is not a linearized theory of gravity. As a consequence, while it still operates according to the rules of quantum field theory, the details are quite different from those for field theories that operate on a fixed classical background space-time. It has considerable successes to its credit, perhaps most notably a quantum theory of spatial geometry, in which quantities such as area and volume are quantized in units of the Planck length, and a calculation of black hole entropy for static and rotating, charged and neutral black holes. But there are also open questions, many of them surrounding the dynamics ("quantum Einstein equations") of the theory.

In contrast to other approaches such as string theory, loop quantum gravity is rather modest in its aims. It is not attempting a grand unification, and hence is not based on an overarching symmetry principle, or some deep reformulation of the rules of quantum field theory. Rather, the goal is to quantize Einstein gravity

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

in four dimensions. While, as we will explain, a certain amount of unification of the description of matter and gravity is achieved, in fact, the question of whether matter fields must have special properties to be consistently coupled to gravity in the framework of loop quantum gravity is one of the important open questions in loop quantum gravity.

Loop quantum gravity is, in its original version, a canonical approach to quantum gravity. Nowadays, a covariant formulation of the theory exists in the so-called *spin foam models*. One of the canonical variables in loop quantum gravity is a connection, and many distinct technical features (such as the "loops" in its name) are directly related to the choice of these variables. Another distinct feature of loop quantum gravity is that no fixed classical geometric structures are used in the construction. New techniques had to be developed for this, and the resulting Hilbert spaces that look very different from those in standard quantum field theory, with excitations of the fields that are one- or two-dimensional. But it has also simplified the theory, since it can be shown that some choices made in the quantum theory are actually uniquely fixed by the requirement of background independence. Furthermore, the requirement of background independence seems to lead to a theory which is built around a very quantum mechanical gravitational "vacuum," a state with degenerate and highly fluctuating geometry. This is exciting, because it means that when working in loop quantum gravity, the deep quantum regime of gravity is "at one's fingertips." However, it also means that to make contact with low-energy physics is a complicated endeavor. The latter problem has attracted a considerable amount of work, but is still not completely solved. Another (related) challenge is to fully understand the implementation of the dynamics. In loop quantum gravity the question of finding quantum states that satisfy "quantum Einstein equations" is reformulated as finding states that are annihilated by the quantum Hamilton constraint. The choices that go into the definition of this constraint are poorly understood in physical terms. Moreover, the constraint should be implemented in an anomaly-free way, but what this entails in practice, and whether existing proposals fulfill this requirement, are still under debate. This is partially due to the lack of physical observables with manageable quantum counterpart, to test the physical implication of the theory.

While these challenges remain, remarkable progress has happened over the last couple of years: the master constraint program has brought new ideas to bear on the implementation of the dynamics [60]. Progress has been made in identifying observables for general relativity that can be used in the canonical quantization [19, 20, 30]. A revision of the vertex amplitudes used in spin foam models has brought them in much more direct contact with loop quantum gravity [25,26]. And, last but not least, in loop quantum cosmology, the application of the quantization strategy of loop quantum gravity to mini-superspace models has become a beautiful and productive laboratory for the ideas of the full theory, in which the quantization program of loop quantum gravity can be tested, and, in many cases, brought to completion [4, 10, 13, 14]. The present review will not cover these developments in any detail, partially because they are ongoing, and partially because there will be separate reviews on group field theory and loop quantum cosmology published alongside the present text. But we hope that it makes for good preparatory reading. In fact, the basic connection between loop quantum gravity and spin foam models is explained in Section 10.3.3, the master constraint program is briefly described in Section 10.3.2, and there are some references to loop quantum cosmology in Section 10.4. Certainly the present review can also not replace the much more complete and detailed reviews that are available. We refer the interested reader in particular to [8, 42, 58].

The structure of this chapter is as follows. In Section 10.2 we explain the classical theory and kinematical quantization underlying loop quantum gravity. Section 10.3 covers the implementation of the Hamilton constraint. In Section 10.4 we consider some physical aspects of the theory: quantized black hole horizons, semiclassical states, and matter propagation. We close with an outlook on open problems and new ideas in Section 10.5.

10.2 Kinematical setup

Loop quantum gravity is a canonical quantization approach to general relativity, thus it is based on a splitting of space-time into time and space, and on a choice of canonical variables. Implicit in the splitting is the assumption that the space-time is globally hyperbolic. Whether topology change can nevertheless be described in the resulting quantum theory is a matter of debate. The choice of canonical variables is characteristic to loop quantum gravity: one of the variables is a connection, and hence the phase space (before implementation of the dynamics) has the same form as that of Yang–Mills theory. As with any canonical formulation of general relativity, the theory has constraints that have to be handled properly both in the classical and in the quantum theory.

The quantization strategy applied in loop quantum gravity is that of Dirac, for the case of first-class constraints. First, a *kinematical* representation of the basic fields by operators on a Hilbert space $\mathcal{H}_{\text{kin}}$ is constructed. In this representation, operators corresponding to the constraints are defined. Then, quantum solutions to the constraints are sought. Such solutions, also called *physical* states, are quantum states that are in the kernel of all the constraints. They form the physical Hilbert space $\mathcal{H}_{\text{phys}}$. Finally, observable quantities are quantized. The corresponding operators should form an algebra $\mathfrak{A}$, and commute with the quantum constraints. Thus $\mathfrak{A}$ leaves $\mathcal{H}_{\text{phys}}$ invariant. The pair $(\mathfrak{A}, \mathcal{H}_{\text{phys}})$ then constitutes the quantum theory of the constrained system in question. Technical aspects of this procedure have to be refined in loop quantum gravity. For example, if the zero eigenvalue is in the continuous part of the spectrum of one of the constraints, the resulting physical space is not part of the Hilbert space but part of its dual. But there are also some fundamental questions about this procedure, such as what guides the choice of the kinematical Hilbert space, and how the quantization and implementation of the constraints are checked. Also, it is notoriously difficult to write down explicit examples of observables for general relativity in the canonical setting, even in the classical theory.

While some of the above questions are not yet answered for loop quantum gravity, the quantum theory is successful in many respects: it includes a fully quantized spatial geometry, and an implementation of the constraints that is anomaly-free at least in a certain sense. In the following, we will give a short, and mostly non-technical introduction to the kinematical aspects of the quantization. The quantization of the Hamilton constraint will be discussed in Section 10.3.

10.2.1 Connection formulation of general relativity

Loop quantum gravity rests on a reformulation of ADM canonical gravity in terms of variables similar to those of Yang–Mills theory. Ashtekar discovered a formulation [2] in terms of a self-dual SL(2,$\mathbb{C}$) connection, and its canonical conjugate, satisfying suitable reality conditions. Loop quantum gravity came to use a formulation in terms of an SU(2) connection [11] for technical reasons. Both of these are actually special cases of a family of formulations depending on several parameters ([39] and literature given therein). We will only consider one of these, the Barbero–Immirzi parameter ι [31]. The covariant description in this case is the Holst action

$$S[e,\omega] = \int \varepsilon^{IJKL} e_I \wedge e_J \wedge F_{IJ}(\omega) + \frac{1}{\iota} e^I \wedge e^J \wedge F_{IJ}(\omega) \qquad (10.1)$$

for an SL(2,$\mathbb{C}$) connection ω and a vierbein e. In the limit $\iota \to \infty$, this is the well-known Palatini action of general relativity. The so-called *Holst term* proportional to ι^{-1} is not a topological term, it depends on the geometry. But, in the absence of fermionic matter, it does not change the equations of motion, as it vanishes identically on shell, due to the Bianchi identity.[1] In the presence of fermions, there are small effects that could in principle be used to distinguish the formulation (10.1) from the Palatini formulation [37].

[1] Actually, instead of adding this term, one can also add the Nieh–Yang term, which *is* topological. The resulting canonical formulation is the same as that with a real Barbero–Immirzi parameter.

The Holst term has a profound effect on the canonical formulation of the theory. A Legendre transform of the Palatini action leads (after solving the second-class constraints) back to the ADM formulation, with spatial metric and exterior curvature as canonical variables. The Legendre transform of (10.1) with *finite* Barbero–Immirzi parameter leads, however, to formulations in which one canonical variable is a connection: for $\iota = \pm i$ the theory has special symmetries and one obtains the Ashtekar formulation [2] in terms of a self-dual SL(2,$\mathbb{C}$) connection. for real ι, and after a partial gauge fixing that gets rid of second-class constraints, one obtains a canonical pair consisting of an SU(2) connection A_a^I and a corresponding canonical momentum E_J^b,

$$\{A_a^I(x), E_J^b(y)\} = 8\pi G\iota\, \delta_a^b \delta_J^I \delta(x, y). \tag{10.2}$$

These fields take values on a spatial slice Σ of the manifold that was chosen in the process of going over to the Hamilton formulation.

There are several constraints on these variables, and the Hamiltonian is a linear combination of constraints. The equations for time evolution are the usual Hamilton equations, and together with the constraint equations they form a set of equations which is completely equivalent to Einstein's equations. The constraints can be written in the following way:

$$G_I = D_a E_I^a \tag{10.3}$$

$$C_a = E_I^b F_{ab}^I \tag{10.4}$$

$$H = \frac{1}{2}\varepsilon^{IJ}{}_K \frac{E_I^a E_J^b}{\sqrt{\det E}} F_{ab}^K - (1+\iota^2)\frac{E_I^a E_J^b}{\sqrt{\det E}} K_{[a}^I K_{b]}^J, \tag{10.5}$$

where D is the covariant derivative induced by A, F is the curvature of A, and K is the extrinsic curvature of Σ in space-time. They have a simple geometric interpretation: G_I generates gauge transformations on phase space. It is also called a *Gauss constraint* to highlight that it is completely analogous to the Gauss-law constraint that shows up in electrodynamics. C_a generates the transformations induced in phase space under diffeomorphisms of Σ. It is therefore also called a *diffeomorphism constraint*. Finally, H generates (when the other constraints hold) the transformations induced in phase space under deformations of (the embedding of) the hypersurface Σ in a timelike direction in space-time. It is also called a *Hamiltonian constraint*, since such deformations can be interpreted as time evolution.

The canonical momentum E has a direct geometric interpretation: it encodes the spatial geometry:

$$|\det q| q^{ab} = E_I^a E_J^b \delta^{IJ}, \tag{10.6}$$

where q_{ab} is the metric induced on Σ by the space-time metric. Thus E is a densitized triad field for q. The interpretation of A is slightly more involved:

$$A_a^I = \Gamma_a^I + \iota K_a^I, \tag{10.7}$$

where Γ is the spin connection related to E.

Matter fields can be added to the canonical description given above. This has to be done with some care, so as not to change the structure of the gravitational sector. For the fermionic sector this requires working with slightly unusual ("half-density") variables [50].

10.2.2 Kinematic representation

The basic variables for the quantization in loop quantum gravity are chosen in such a way as to make their transformation behavior under SU(2) and spatial diffeomorphisms as simple and transparent as possible. The obvious reason behind this goal is that one wants to simplify the solution of the constraints as much as possible. Early ideas about this go back to Rovelli and Smolin [43]. We follow here [5]. An obvious choice for the connection A is its holonomies

$$h_\alpha[A] = \mathcal{P}\exp\int_\alpha A, \tag{10.8}$$

or more generally, functions of such holonomies

$$f[A] \equiv f(h_{\alpha_1}[A], h_{\alpha_2}[A], \ldots, h_{\alpha_n}[A]), \tag{10.9}$$

for a finite number of paths $\alpha_1, \ldots, \alpha_n$. Such functionals are also called *cylindrical functions*.

For the field E a natural functional is its flux through surfaces S:

$$E[S, f] = \int_S *E_I f^I, \tag{10.10}$$

where f is a function taking values in SU(2)* and $*E$ is the 2-form $E^a \varepsilon_{abc} \mathrm{d}x^b \wedge \mathrm{d}x^c$.

To quantize cylindrical functions and fluxes, one is seeking a representation of the following algebraic relations on a Hilbert space:

$$\begin{aligned}
f_1 \cdot f_2[A] &= f_1[A] f_2[A] \\
[f, E_{S,r}] &= 8\pi \iota l_P^2 X_{S,r}[f] \\
[f, [E_{S_1,r_1}, E_{S_2,r_2}]] &= (8\pi \iota l_P^2)^2 [X_{S_1,r_1}, X_{S_2,r_2}][f] \\
&\cdots \\
(E_{S,r})^* = E_{S,\bar{r}}, &\qquad (f[A])^* = \overline{f}[A].
\end{aligned} \tag{10.11}$$

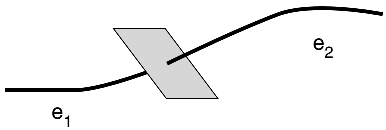


Figure 10.1 Illustration of the edges involved in (10.21).

Here, X is a certain derivation on the space of cylindrical functions. As an example, consider the case of a surface S that is intersected transversally by a path e, splitting it into a part e_1 incoming to, and a part e_2 outgoing from, the surface (see Figure 10.1). Then (with a certain orientation of the surface assumed)

$$X_{S,r}\overset{j}{\pi}(h_e) = \sum_i r_i(p)\overset{j}{\pi}(h_{e_1}\tau_i h_{e_2}). \tag{10.12}$$

The commutators between cylindrical functions and fluxes come from the Poisson relations (10.2). It is somewhat surprising to see that there are also non-trivial commutators between fluxes. These are required to turn the algebra of fluxes and cylindrical functions into a Lie algebra, a structure that has representations in terms of operators on Hilbert spaces.

Loop quantum gravity employs a specific representation of (10.11) on a Hilbert space $\mathcal{H}_{\text{kin}}$. A basis for this Hilbert space is given by the so-called generalized spin networks. Such a network is by definition an oriented graph γ embedded in Σ, together with a labeling of the edges and vertices of that graph. The edges are labeled by irreducible representations of SU(2). A vertex carries elements of the dual of the tensor product of all representations on the edges that are incoming to or outgoing from the vertex as a label (see Figure 10.2). A generalized spin network represents a way of constructing a cylindrical functional. To obtain its value on a given connection, one computes the holonomies along the edges of the graph in the representations given by the edge labels, and contracts these via the labels of the vertices. And, vice versa, any cylindrical function can be written as a (possibly infinite) linear combination of generalized spin networks.

An inner product is then defined on the span of these generalized spin networks by postulating that they are orthogonal to each other, and by specifying their norm in terms of the labels of the graph. This inner product is completely invariant under the action of the diffeomorphisms of Σ.

A representation of (10.11) is given on generalized spin networks by using the fact that they can be viewed as cylindrical functionals. The cylindrical functionals can thus be represented as multiplication operators, the fluxes by derivations

$$(f\psi)[A] = f[A]\psi[A], \qquad (E_{S,r}\psi)[A] = (X_{S,r}\psi)[A]. \tag{10.13}$$

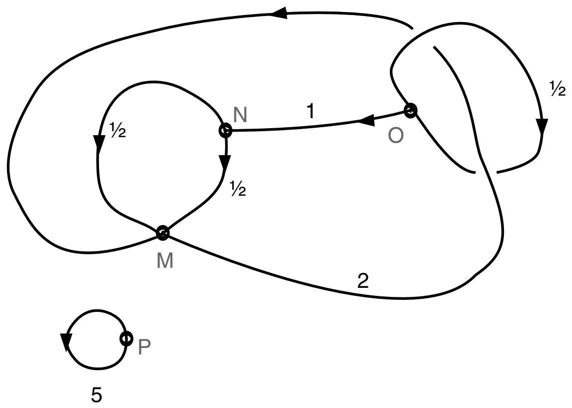


Figure 10.2 A generalized spin network. The labels M, N, O, P can be thought of as suitably dimensional tensors.

As with the inner product, these definitions do not make use of any background structure, such as a classical metric. They are thus covariant under the action of the diffeomorphisms of Σ. Moreover, there is a state in $\mathcal{H}_{\mathrm{kin}}$ that is even *invariant* under the action of those diffeomorphisms. This state is the empty (i.e. without any edges or vertices) generalized spin network. Moreover, any state in the kinematical Hilbert space can be approximated by applying linear combinations of products of the basic operators to this diffeomorphism-invariant state. In mathematical language, this state is therefore cyclic. It, together with the algebraic relations between the basic variables, completely encodes the structure of the kinematical representation.

We note also that the kinematical representation has the following peculiar properties:

(1) The diffeomorphisms ϕ of Σ are represented on $\mathcal{H}_{\mathrm{kin}}$ by unitary operators U_ϕ. This follows from what we have already said about their action. But generators for these unitary operators do not exist. If $\phi(t)$ is a one parameter family of diffeomorphisms, with $\phi(0) = \mathbb{I}$, then

$$\frac{1}{i}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 U_{\phi(t)} \tag{10.14}$$

does not exist, in any sense, as a well-defined operator.

(2) We have seen that the holonomies $h_e[A]$ exist as matrices of operators. But neither can one obtain from them an operator for the curvature F, nor for the connection A itself. The limits

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^2}\left(h_{\alpha_\varepsilon} - \mathbb{I}\right), \qquad \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(h_{\beta_\varepsilon} - \mathbb{I}\right) \tag{10.15}$$

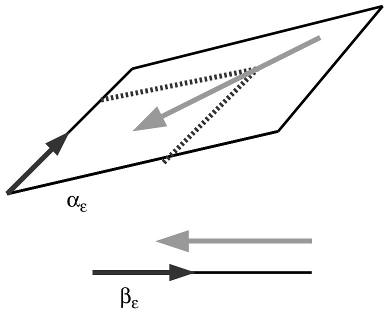


Figure 10.3 The loop α_ε and the edge β_ε of (10.15).

do not exist in any sense as well-defined operators on $\mathcal{H}_{\text{kin}}$. α_ε is here a plaquette loop with (coordinate) side length ε, and β_ε is an open line with (coordinate) side length ε (see Figure 10.3).

It may appear that a lot of choices have been made in the definition of $\mathcal{H}_{\text{kin}}$ and the representation of the basic variables on it. But this is not the case. The following uniqueness theorem can be proven [28, 35].

Theorem 10.1 *Any representation of the algebraic relations* (10.11) *that contains a diffeomorphism-invariant cyclic vector is equivalent to the one on $\mathcal{H}_{kin}$ described above.*

Diffeomorphism invariance should be seen here as a requirement dictated by the philosophy of loop quantum gravity (no use of geometric background structure), as well as by simplicity (implementation of the diffeomorphism constraint consists precisely in throwing out any non-diffeomorphism-invariant information). While cyclicity would be a requirement on the physical sector, here it is only a natural simplification.

10.2.3 Geometric operators

It is possible to quantize areas and volumes with respect to the geometry on Σ on the Hilbert space $\mathcal{H}_{\text{kin}}$ [6, 7, 44]. Since the quantum Einstein equations, in the form of the constraints, have not yet been taken into account, the physical implications of the results have to be considered with substantial care [24, 41]. There are, however, situations in which such quantities are observables, in the sense that they commute with the constraints. This is for example the case with the area of a black hole horizon as considered in Section 10.4.1 below. In such cases the results that we are going to present have clear physical significance.

We consider the case of areas. Let S be a surface in Σ. When the field E is pulled back to Σ one obtains a vector-valued 2-form. The norm of this 2-form is directly related to the area [40]:

$$A_S = \int_S |E|. \tag{10.16}$$

This formula can be used as a starting point for quantization. Regularizing in terms of fluxes in the form of (10.10), substituting operators, and taking the regulator away leads to a well-defined, simple operator $\widehat{A}_S$. Its action on states with just a single edge is especially simple: if edge and surface do not intersect, the state is annihilated. If they do intersect once, one obtains

$$\widehat{A}_S \,\mathrm{Tr}[\pi_j(h_\alpha[A])] = 8\pi\iota l_P^2 \sqrt{j(j+1)}\,\mathrm{Tr}[\pi_j(h_\alpha[A])]. \tag{10.17}$$

Thus these states are eigenstates of area, with the eigenvalue given as the square root of the eigenvalue of the SU(2) Casimir in the representation given on the edge. A slightly more complicated action is obtained in the case of several intersections, and in particular if a vertex of the generalized spin network lies within the surface. Nevertheless, the area operator can be completely diagonalized. It turns out that the spectrum is discrete. As is seen in (10.17), the scale is set by Planck area l_P^2 (Figure 10.4). The eigenvalue density increases exponentially with area. A similar procedure leads to an operator for volumes of subregions in Σ. This operator is substantially more complicated. Unlike the area operator, the action of which is purely in terms of the representation label of the edges, the volume operator acts on the vertices, by changing the maps that label them ("recoupling"). In fact, there are two slightly different versions of the volume operator [7, 44], differing in the way the tangent space structure of a vertex is taken into account. In either case, the

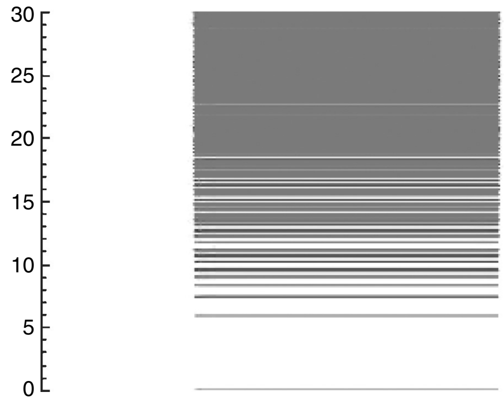


Figure 10.4 The lowest part of the area spectrum of loop quantum gravity, in units of l_P^2.

spectrum is discrete, but not explicitly known. A beautiful computer analysis of the lowest part of the spectrum can be found in [16, 17].

10.2.4 Gauge-invariant states, spin networks

The simplest of the constraints (10.3) to implement is the Gauss constraint: $G_I = D_a E^a_I$. It can easily be checked that classically, it generates SU(2) transformations, which act on holonomies as

$$h_e[A] \mapsto g(s(e))h_e g(t(e))^{-1} \tag{10.18}$$

with $g(x)$ the gauge transformation, and $s(e), t(e)$ the beginning and endpoint of e. Thus, there are two ways to implement this constraint. One can regularize the expression for G_I in terms of holonomies and fluxes, which have well-defined quantization, quantize the regularized expression, and remove the regulator, hoping to obtain a well-defined constraint operator in the limit. If successful, one can then determine the kernel of the quantum constraint. Or one can declare that all states in $\mathcal{H}_{\text{kin}}$ that are invariant under gauge transformations (10.18) are solutions to the constraint. Both strategies are viable, and lead to exactly the same result: The solution space $\mathcal{H}_{\text{gauge}}$ is a proper subspace of $\mathcal{H}_{\text{kin}}$. An orthonormal basis is given by the so-called *spin networks* [45]. These are special cases of the generalized spin networks, in that the linear maps labeling the vertices are intertwining operators

$$I_v : \bigotimes_{e \text{ incoming}} \pi_{j(e)} \longrightarrow \bigotimes_{e \text{ outgoing}} \pi_{j(e)}, \qquad I_v \pi_{\text{incoming}}(g) = \pi_{\text{outgoing}}(g) I_v \tag{10.19}$$

mapping the tensor product of the representations on the incoming edges to the tensor product of the representations on the outgoing edges. The contraction of the holonomies with these intertwiners guarantees that the resulting states are invariant under gauge transformations.

10.2.5 Diffeomorphism-invariant states

The diffeomorphism constraint $C_a = E^b_I F^I_{ab}$ has not been quantized directly. One reason is that curvature cannot be quantized on $\mathcal{H}_{\text{kin}}$ but one can see even on more general grounds that a quantization of C_a must run into difficulties: Classically, this constraint generates the diffeomorphisms of Σ, and one expects the same of its quantum counterpart. Otherwise one would have produced an anomalous implementation of the constraint, with possibly disastrous consequences for the theory. But the diffeomorphisms ϕ of Σ already act on $\mathcal{H}_{\text{kin}}$, through unitary operators U_ϕ.

These operators are however, not strongly continuous in the diffeomorphisms (see (10.14)), in other words, they have no self-adjoint generators. Thus C_a cannot be directly quantized without generating anomalies. But this is not a problem, as we know what the gauge transformations generated by C_a are, and because they are acting in a simple manner on $\mathcal{H}_{\text{kin}}$. All one has to do is find states that are invariant under the action of the diffeomorphisms U_ϕ.

The action of the diffeomorphisms on cylindrical functions consists in moving the underlying graph:

$$U_\varphi \psi_\gamma = \psi_{\varphi(\gamma)}. \tag{10.20}$$

Therefore, the only invariant state in $\mathcal{H}_{\text{gauge}}$ is the empty spin network. Rather than in $\mathcal{H}_{\text{diff}}$, the rest of the invariant states lie in the dual of $\mathcal{H}_{\text{diff}}$. They can be found by group averaging. This procedure assigns to a state $\psi \in \mathcal{H}_{\text{gauge}}$ a diffeomorphism-invariant functional $\Gamma\psi$. The idea is

$$(\Gamma\psi)(\phi) = (\text{Vol}(\text{Diff}))^{-1} \int_{\text{Diff}} D\varphi \, \langle U_\varphi \psi \,|\, \phi \rangle_{\mathcal{H}_{\text{kin}}}. \tag{10.21}$$

This is still formal. To make this work, the integration over the diffeomorphism group, and the division by its volume, have to be made sense of. These tasks would be hopeless, were it not for the unusual properties of the scalar product on $\mathcal{H}_{\text{kin}}$. In fact, the correct notion in this context of the integral over diffeomorphisms is that of a sum! A careful examination leads to the formula [8, 9]

$$(\Gamma\psi_\gamma)(\phi) = \sum_{\varphi_1 \in \text{Diff}/\text{Diff}_\gamma} \frac{1}{|\text{GS}_\gamma|} \sum_{\varphi_2 \in \text{GS}_\gamma} \langle \varphi_1 * \varphi_2 * \psi_\gamma \,|\, \phi \rangle. \tag{10.22}$$

Here, Diff_γ is the subgroup of diffeomorphisms mapping γ onto itself, and TDiff_γ the subgroup of Diff which is the identity on γ. The quotient $\text{GS}_\gamma := \text{Diff}_\gamma/\text{TDiff}_\gamma$ is called the set of *graph symmetries*. It can be checked that this definition really gives diffeomorphism-invariant functionals over $\mathcal{H}_{\text{gauge}}$. An inner product can also be defined on these functionals, using (10.22). Thus one obtains a Hilbert space $\mathcal{H}_{\text{diff}}$ of gauge and diffeomorphism-invariant quantum states.

It is sometimes stated that diffeomorphism-invariant spin network states are labeled by equivalence classes of spin networks under diffeomorphisms. This is a nice intuitive picture, but one has to be careful with it: the effects of (10.22) can be quite subtle. For example, the map Γ has a large kernel. Some spin networks, such as the "hourglass" (see Figure 10.5), are mapped to zero.

Diffeomorphism-invariant quantities can give rise to well-defined operators on $\mathcal{H}_{\text{diff}}$. An example is the total volume V_Σ of Σ. The corresponding operator on $\mathcal{H}_{\text{kin}}$ extends to $\mathcal{H}_{\text{diff}}$, thus one obtains a well-defined notion of quantum volume. Areas

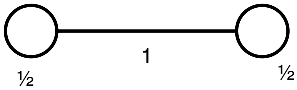


Figure 10.5 The hourglass spin network gets mapped to zero under group averaging with respect to the diffeomorphism group.

of surfaces and volumes of subregions of Σ can similarly be quantized, provided surfaces and regions can be defined in a diffeomorphism-invariant fashion, for example by using a matter field as reference system.

10.3 The Hamilton constraint

The most complicated constraint is the Hamilton constraint

$$H = \underbrace{\frac{1}{2}\varepsilon^{IJK}\frac{E^a_I E^b_J}{\sqrt{\det E}} F_{abK}}_{H_E} - (1+\iota^2)\frac{E^a_I E^b_J}{\sqrt{\det E}} K^I_{[a} K^J_{b]}. \tag{10.23}$$

Here we have already denoted by H_E the so-called *Euclidean part* of the constraint, which we will need later. The quantization of the Hamilton constraint poses several difficulties. On the one hand, its classical action is very complicated on the basic fields A and E. Therefore methods based on a geometric interpretation, such as were used to find solutions to the diffeomorphism constraint, are not available. Its functional form, on the other hand, makes it hard to quantize in terms of the basic fields because it contains (a) the inverse volume element, and (b) the curvature of A. (a) is problematic because large classes of states in $\mathcal{H}_{\text{diff}}$ have zero volume, thus its inverse tends to be ill defined. There have to be subtle cancellations between the inverse volume and other parts of the constraint for the whole to be well defined. (b) is problematic, because its curvature cannot be quantized in a simple way, at least on $\mathcal{H}_{\text{kin}}$, due to the nature of the inner product. It is thus very remarkable that Thiemann [51–55] proposed a family of well-defined Hamiltonian constraints, and partially analyzed the solution spaces. We cannot describe his construction with all details, but we will briefly discuss the most important ideas.

10.3.1 Thiemann's quantization

The first ingredient in the quantization is the observation that one can absorb the inverse volume element in the Hamiltonian constraint into a Poisson bracket between the connection and the volume:

$$\varepsilon^{IJK}\varepsilon_{abc}\frac{E^a_I E^b_J}{\sqrt{\det E}} = \frac{1}{4\iota}\{A^K_c(x), V_\Sigma\}. \tag{10.24}$$

Here V_Σ is the volume of the spatial slice. The Poisson bracket can be quantized as a commutator

$$\{\ldots,\ldots\} \longrightarrow \frac{1}{i\hbar}[\ldots,\ldots], \tag{10.25}$$

and the volume has a well-understood quantization as we have discussed before. A similar trick can also be used to quantize the extrinsic curvature appearing in the Hamilton constraint. Thiemann found that

$$K^I_a E^a_I(x) = \{H_E(x), V_\Sigma\}, \tag{10.26}$$

which can be used to quantize the full constraint, once the Euclidean part H_E has been quantized.

The second important idea is that solutions to all the constraints must, in particular, be invariant under spatial diffeomorphisms. Thus it is possible to define operators for curvature as a limit of holonomy around shrinking loops. While such limits are ill defined when acting on kinematical states, they can be well defined on states in $\mathcal{H}_{\text{diff}}$. Indeed, Thiemann is able to give a regulated definition for the constraints, which is such that when evaluated on states in $\mathcal{H}_{\text{diff}}$, it becomes independent of the regulator, once it is small enough. The Poisson bracket involving A can be approximated as

$$\varepsilon\vec{e}^a\{A_a(x), V_\Sigma\} \approx \{\int_e A, V_\Sigma\} \approx -h_e^{-1}\{h_e, V_\Sigma\}, \tag{10.27}$$

where e is a curve emanating in x, $\vec{e}$ is its tangent in x, and ε its coordinate length. Curvature is treated as in lattice gauge theory

$$\varepsilon^2 F_{ab}(x)\mathrm{d}\sigma^{ab} \approx \int_S F \approx h_{\partial S} - \mathbb{I}, \tag{10.28}$$

where $\mathrm{d}\sigma$ is the area element of the surface S in x and ε^2 its coordinate area. To use the formulas (10.27), (10.28), one needs to choose curves and surfaces. For the Hamilton constraint, these are made to depend on the graph that the state acted on, and they are assigned in a diffeomorphism-covariant fashion. This still leaves large ambiguities when the operators acts on states in $\mathcal{H}_{\text{kin}}$, but most of them go away, when acting on $\mathcal{H}_{\text{diff}}$: only their diffeomorphism-invariant properties matter.

Using these ideas, one can define Hamilton constraint operators on $\mathcal{H}_{\text{kin}}$, and by duality on $\mathcal{H}_{\text{diff}}$. The operators have the following properties:

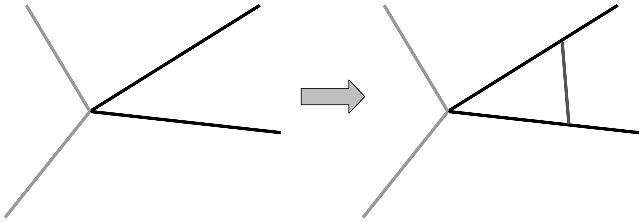


Figure 10.6 The Hamilton operator creates new edges between edges incident in a common vertex.

- The action of the constraints is local around the vertices:

$$\widehat{H}(N)\psi_\gamma = \sum_{v\in V(\gamma)} N(v)\widehat{H}(v)\psi_\gamma, \tag{10.29}$$

  where ψ is a cylindrical function based on the graph γ, the sum is over the vertices of γ, and $\widehat{H}(v)$ is an operator that acts only at, and in the immediate vicinity of, the vertex v.
- They create new edges (see Figure 10.6). This is because of the use of (10.28): the surfaces chosen to regulate the curvature are such that one of the edges bounding them is not part of the graph of the state acted on.
- They have a non-trivial kernel.
- They are anomaly-free in a certain sense: the commutator of two constraints vanishes on states in $\mathcal{H}_{\text{diff}}$. See the discussion below.
- There are several ambiguities in the definition of the constraints. One is for example the SU(2) representation chosen in the regularization process (in (10.27), (10.28), for example, we worked with holonomies in the defining representation, but other irreducible representations could be used as well). But there are also ambiguities pertaining to the creation of new edges, and ambiguities in the application of equations like (10.24) that are harder to parametrize.

We do not know whether the quantization proposed by Thiemann is the right one. One important test for the quantization of constraints is whether they satisfy the relations that are expected from the classical Dirac algebra. The commutator of two Hamilton constraints is expected to be a diffeomorphism constraint. Thiemann's quantization is anomaly-free in the sense that the commutator of two Hamilton constraints vanishes when evaluated on a diffeomorphism-invariant state [54]. But it was found that the same holds when the commutator is evaluated on a much larger set of states, for which the diffeomorphism constraint is not expected to vanish [29, 34]. Also there are several ambiguities in Thiemann's quantization. The meaning of these is largely unclear, but some have been investigated [12, 36]. Ultimately, the questions surrounding the quantization of the Hamiltonian should be answered by physical considerations, for example by checking the classical limit of the theory, or by other predictions that the theory makes. One situation in which such questions can be posed and answered is loop quantum cosmology, and we expect important input from the findings there. For a technical solution to some of the problems with the constraint algebra see the next section, on the master constraint program. Also, substantial progress concerning the dynamics of the theory has been made in the spin foam approach, and we hope it will shed light on the issues here (see Section 10.3.3 and the contribution by Oriti to this volume). In summary: while many aspects need more study, there is no doubt that Thiemann's work on the Hamiltonian contains at least part of the solution of the problem of dynamics in LQG.

10.3.2 The master constraint approach

As we have pointed out above, the question of whether Thiemann's proposal for the quantization of the Hamilton constraint is anomaly-free is a question that is not settled. In fact, the Poisson relations between two Hamiltonian constraints are very complicated and involve the phase space point. The resulting algebra is thus not a Lie algebra, and it is unclear what a representation of it should look like. In particular, one expects some quantum deformations of the structure to occur, but just what constitutes a (harmless) deformation, and what a (harmful) anomaly, is not clear. These difficulties prompted the proposal of the master constraint program [60]. At its core, the proposal is to replace implementation of the infinite-dimensional algebra of constraints with the implementation of one master constraint. In the case of the Hamiltonian constraint, the proposal is to go over to the quantity

$$\mathbf{M} = \int_S \frac{(H(x))^2}{\sqrt{\det q}(x)}. \tag{10.30}$$

It can be argued on general grounds, and checked in examples, that the kernel of the quantization of such a master constraint $\mathbf{M}$ is the same as the joint kernel of the individual constraints constituting the master constraint. It is obvious that in this way questions about the constraint algebra can be alleviated. In the case of loop quantum gravity, one can even add squared diffeomorphism and Gauss constraints to the master constraint above, thus reducing the considerations to only one constraint altogether. The master constraint is then much more complicated than the original constraints, but quantization can be attempted with similar techniques as were used for the Hamilton constraint, described above.

The master constraint method has been tested extensively (see for example [21–23]), and appears to afford a large simplification in many cases. In eliminating the constraint algebra, it does however do away with an important check for the correctness of the quantization. If there are other good ways to check this correctness, this is no problem, but in cases – such as at the present moment the quantization of the constraints in loop quantum gravity – in which no other good means of checking the quantization exist, its application is not without danger.

10.3.3 Physical inner product, and the link to spin foam models

For physical applications it is not merely the physical states that are important. To compute amplitudes and expectation values one needs an inner product on these states. In theory, this inner product is obtained from the constraints themselves. If their joint kernel is contained in the kinematical Hilbert space, the inner product on that space simply induces one on $\mathcal{H}_{\text{phys}}$. If zero is in the continuous spectrum of some of the constraints, there are still mathematical theorems that guarantee the existence of an inner product, but it can be extremely hard to compute in practice.

We now want to describe a formal series expansion of the inner product on $\mathcal{H}_{\text{phys}}$ due to Rovelli and Reisenberger [38]. Since it is formal, it may not necessarily be useful to calculate the inner product exactly, but it is hugely important because it makes contact with approaches to quantum gravity that are starting from discretizations of the path integral of gravity, so-called *spin foam models*. With that, it brings back into loop quantum gravity an intuitive image of time evolution. This is very important, even if its physical merits are still under debate.

The series expansion is obtained by considering the projector P_{phys} on the Hilbert space $\mathcal{H}_{\text{phys}}$ of physical states: each of the Hamilton constraints comes with a projector on its kernel. This may be a genuine projection operator, or a linear map into the dual of $\mathcal{H}_{\text{diff}}$. It can formally be written as $P^{(x)}_{\text{phys}} = \delta(\widehat{H}(x))$. The projection onto the solution space of the Hamilton constraints is the product of all these projectors.

The physical inner product between the (physical part of) spin networks ψ and ψ' can be expressed in terms of the projector as

$$\langle P_{\text{phys}}\psi \mid P_{\text{phys}}\psi'\rangle_{\text{phys}} = \langle \psi \mid P_{\text{phys}} \ldots\rangle_{\mathcal{H}_{\text{diff}}}. \tag{10.31}$$

To obtain the series expansion, one writes the delta functions as functional integration over the lapse, and expands the exponential:

$$\begin{aligned} P_{\text{phys}} &= \prod_{x\in\Sigma} \delta(\widehat{H}(x)) \\ &= \int DN \exp i \int N(x)\widehat{H}(x)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= 1 + i \int DN \int N(x)\widehat{H}(x)\,\mathrm{d}x \\
&\quad - \frac{1}{2}\int DN \iint N(x)N(x')\widehat{H}(x)\widehat{H}(x')\,\mathrm{d}x\,\mathrm{d}x' \\
&\quad + \ldots . \qquad (10.32)
\end{aligned}
$$

One sees that the expansion parameter is the number of Hamilton constraints in the expression. It was shown in [38] how the path integrals over N can be defined. If one plugs this expansion into (10.31), one obtains an expansion of the physical inner product in terms of the product on $\mathcal{H}_{\text{diff}}$. The Hamilton constraints will create and destroy edges. In fact, each term in (10.32) will give rise to multiple terms in the expansion of the inner product, as each of the Hamilton constraints can, because of the integration over Σ, act at any of the vertices. In the end, many of the terms will give zero, however, because the scalar product on $\mathcal{H}_{\text{diff}}$ is non-zero only if the graphs underlying the states are equivalent under diffeomorphisms. This means that each of the non-zero terms can be labeled by a diagram that depicts a discrete cobordism, or history, connecting the two graphs involved in the product. For an illustration, see Figure 10.7. In such a diagram, surfaces represent the evolution of the edges of the spin networks, these meet in lines, which represent the evolution of the spin network vertices. When a new edge is created by the action of a constraint, this is shown in the diagram as a vertex. Such diagrams, together with the labeling of the surfaces with representations, and the edges with intertwiners, is called a spin foam. Each of these spin foams is assigned, by the action of the Hamilton constraints and the inner product on $\mathcal{H}_{\text{diff}}$, a number (amplitude). This is very interesting for several reasons:

- The analogy to Feynman diagrams is striking. In both cases, an evolution operator is expanded into a series of terms labeled by topological objects with group representations as labels.

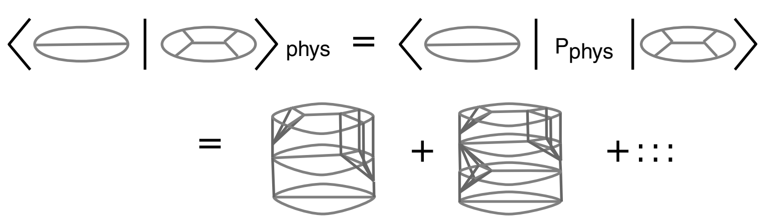


Figure 10.7 An illustration of the spin foams occurring in the expansion of the physical inner product of two specific spin networks. Spin labels are not shown for simplicity.

- Spin foam models have been obtained independently, from discretizations of the action of general relativity.
- Solving the Hamilton constraint means implementing the dynamics of loop quantum gravity, but no notion of evolution is apparent in solutions, at least superficially. The above expansion brings back a picture of state evolution (although one must be cautious with simple physical interpretations in terms of geometry evolving in some specific time).

While the connection between loop quantum gravity and spin foam models described above is very convincing in abstract terms, when one compares the models one gets from using, for example, Thiemann's constraint with spin foam models obtained independently, there are big technical differences – starting from the notion of the graphs involved (embedded vs abstract) and not ending with the groups involved. Some of this is changing, however. For the interesting new perspectives that result, we refer the reader to [18, 25].

10.4 Applications

Now, after the complete formalism of loop quantum gravity has been laid out, we can come to some applications. However, it is presently impossible to solve the constraints in all generality, and investigate their physical properties. This is due, on the one hand, to the difficulties with the implementation of the Hamilton constraints (see Section 10.3), and on the other hand to the absence of useful observables that can be quantized, and used to investigate physical states. As an example, we remark that the question of whether a space-time contains black holes or not is well defined, and can in principle be answered in terms of initial values on a spatial slice Σ. But to do this in practice is a very difficult task in the classical theory, and clearly beyond our abilities in the quantum theory. Therefore, simplifying assumptions and approximations have to be made. We will report here on studies on the quantum theory of a horizon of a black hole (in which the existence of a null-boundary and some of its symmetries are presupposed), and on some approximations, called *semiclassical states*, to physical states and their application to the calculation of matter propagators. Another area with important physical applications is loop quantum cosmology, in which the techniques (and in some cases, results) of loop quantum gravity are applied to mini-superspace models. A separate review is covering this area in detail. Finally we mention the research in spin foam models which has led to a program to determine the graviton propagator.

10.4.1 Black holes

Black holes are fascinating objects predicted by general relativity. They even point beyond the classical theory, because of the singularities within, and because of the intriguing phenomenon of black hole thermodynamics [61]. Therefore they are a tempting subject of investigation in any theory of quantum gravity. Loop quantum gravity was able to successfully describe black hole horizons in the quantum theory. Within this description, it is possible to identify degrees of freedom that carry the black hole entropy, and prove, for a large class of black holes, the Bekenstein–Hawking area law. The classic treatment is by Ashtekar, Baez, and Krasnov [3], while first ideas go back to Krasnov and Rovelli [33]. Recently, a partial gauge fixing that had been employed in [3] was found unnecessary [27]. There are some quantitative differences with the original treatment, but the qualitative picture stays exactly the same. In our description, we will follow [27].

The loop quantum gravity calculation does not start from solutions of the full theory. Rather, it quantizes gravity on a manifold with boundary Δ. In the simplest case, the boundary is assumed to be null, with topology $\mathbb{R}\times S^3$. Again, there are fields A and E on a manifold Σ, but now Σ has a boundary H. The boundary Δ is now required to be an *isolated horizon*, a quasi-local substitute for an event horizon. This imposes boundary conditions on the fields A and E at H,

$$*E = -\frac{a_H}{\pi(1-\iota^2)}F(A). \tag{10.33}$$

a_H denotes the area of the horizon H. Furthermore, the symplectic structure acquires a surface term. The latter suggests, together with some technical aspects of the kinematical Hilbert space used in loop quantum gravity, to quantize the fields on the horizon separately from the bulk fields. The latter are quantized in the way described in Section 10.2. The only new aspect is that now edges of a spin network can end on the horizon. Such ends of spin network edges are described by quantum numbers $m_p \in \{-j_p, -j_p+1, \ldots, j_p-1, j_p\}$, where j_p is the representation label of the edge ending on the horizon, and p is a label for the endpoint ("puncture"). The quantum number represents the eigenvalue of the component of E normal to the horizon at the puncture.

The boundary term in the symplectic structure is that of a SU(2) Chern–Simons theory with level

$$k = \frac{a_H}{2\pi\iota(1-\iota^2)l_P^2}, \tag{10.34}$$

and punctures where spin network edges of the bulk theory end on the surface. The quantized Chern–Simons connection is flat, locally, but there are degrees of freedom at the punctures. These are – roughly speaking – described by quantum numbers

s_p, m'_p, where the former is a half-integer, and $m'_p \in \{-s_p, -s_p+1, \ldots, s_p-1, s_p\}$. There is a constraint on the set of m'_p's coming from the fact that H is a sphere, and hence a loop going around all the punctures is contractible, and the corresponding holonomy must hence be trivial. The Hilbert space is equivalent to a subspace of the singlet component of the tensor product $\pi_{s_1} \otimes \pi_{s_2} \otimes \ldots$ ranging over all punctures. The boundary condition (10.33) can be quantized to yield an operator equation. The solutions are tensor products of bulk and boundary states in which the quantum numbers (s_p, m'_p) and (j_p, m_p) are equal to each other at each puncture.

Now, if one fixes the quantum area of the black hole to be a, this bounds the number of punctures and the spins (j_p) labeling the representations. It becomes a rather complicated combinatorial problem to determine the number $N(a)$ of quantum states with area a that satisfy the quantum boundary conditions. It was solved in [1]. It turns out that

$$S(a) := \ln(N(a)) = \frac{\iota}{\iota_{ENP}} \frac{a}{4\pi l_P^2} - \frac{3}{2} \ln \frac{a}{l_P^2} + O(a^0) \tag{10.35}$$

as long as $\iota \leq \sqrt{3}$. Here, ι_{ENP} is the constant that solves the equation

$$1 = \sum_{k=1}^{\infty} (k+1) \exp\left(-\frac{1}{2}\iota_{ENP}\sqrt{k(k+2)}\right). \tag{10.36}$$

One finds $\iota_{ENP} \approx 0.274$. One thus obtains the Bekenstein–Hawking area law upon setting $\iota = \iota_{ENP}$.

10.4.2 Semiclassical states and matter propagation

As we have seen before, the trivial spin network is a diffeomorphism-invariant cyclic vector, in a sense, the *vacuum* of loop quantum gravity. This state has the spatial geometry completely degenerate, and the connection field A maximally fluctuating. It is a solution to all the constraints, yet it does not look at all like a classical space-time. Therefore one needs to look for states that behave more like a classical space-time geometry. While it would be desirable to find such states that at the same time also satisfy all the constraints, this has not been achieved so far in the full theory (the situation is much better in loop quantum cosmology, though – see for example [10]). Rather, one is settling for states that approximate a given classical metric, and at the same time are approximate solutions to the constraints. Such states have come to be called *semiclassical states*. They are useful for studying the classical limit of the theory, as well as for attempting predictions, and as a starting point for perturbation theory.

One particular class of states that has been studied is using coherent states for the group SU(2) [46,56,57,59]. To understand these states, it is useful to remember the coherent states for the harmoninc oscillator:

$$z := \frac{1}{\sqrt{2}}\left(\frac{1}{\sigma}x_0 + i\frac{\sigma}{\hbar}p_0\right), \qquad \psi_z^\sigma(x) \sim \left[e^{-\sigma^2\Delta}\delta_w\right]_{w\to z}(x). \tag{10.37}$$

Thus, coherent states can be viewed as analytic continuations of the heat kernel. This viewpoint makes generalization to a compact Lie group G possible:

$$\psi_h^t(g) := \left[\exp\left(-t\Delta^G\right)\delta_w^G\right]_{w\to c}(g) \equiv \left[\sum_\pi d_\pi e^{-t\lambda_\pi}\chi_\pi(gw^{-1})\right]_{w\to c}(g). \tag{10.38}$$

These states are of minimal uncertainty in a specific sense, and are moreover sharply peaked at a point of T^*G encoded in h. These states can be used in loop quantum gravity. The idea is to use a random graph γ which is isotropic and homogeneous on large scales, together with a cell complex dual to γ. In particular, there will be a face S_e dual to each edge e of γ. Now, given a classical phase space point (A, E), one defines

$$c_e := \exp\left[i\tau_j \int_{S_e} \star E^j\right] h_e(A), \tag{10.39}$$

and then

$$\Psi^t_{\gamma,(A,E)} := \bigotimes_{e\in\gamma} \Psi^t_{c_e} \qquad \in \mathrm{cyl}_\gamma \subset L^2(\overline{\mathcal{A}}, d\mu_{AL}). \tag{10.40}$$

These states satisfy the Hamilton constraint weakly, in the sense that the expectation value of the constraint vanishes and they are strongly peaked at a classical solution.

Such states can be used to approximately compute matter dispersion relations [48,49]. The situation studied is that of a matter test field propagating on quantum space-time. Two scenarios have been investigated:

(1) The field is coupled to the expectation values (in a semiclassical state) in the gravity sector with semiclassical state.
(2) The geometry is chosen as the “typical result” of a measurement in the gravity sector that has been in a semiclassical state.

Either case results in a coupling of the matter field to a fluctuating, discrete spatial geometry. In a (1+1)-dimensional toy model for a scalar field, the dispersion relation has been explicitly calculated [47]:

$$\omega(k) = c^2 + \ell^2 k^4 + O(k^6), \tag{10.41}$$

with

$$
\begin{aligned}
c^2 &= \lim_{N\to\infty} \frac{\langle l\rangle^2}{\langle l^2\rangle} = \frac{1}{1+\frac{d^2}{l^2}} \\
\ell^2 &= \lim_{N\to\infty} \left(\frac{1}{N^2}\frac{\langle l\rangle^4}{\langle l^2\rangle^3} \sum_{i<j} c_{ij} l_i^2 l_j^2 - \frac{N^2}{12}\frac{\langle l\rangle^4}{\langle l^2\rangle} \right) \\
&= -\frac{1}{12} l^2 \frac{1}{1+\frac{d^2}{l^2}}.
\end{aligned} \tag{10.42}
$$

N is here the size of a lattice with periodic boundary conditions, and $\langle\cdot\rangle$ denotes averages over the random lattice. l is the average effective lattice spacing, and d is a measure of the fluctuation in the latter. The phase velocity c depends on the details of state and graph, and may do so differently for different fields. This opens the door to obtaining severe constraints on the theory from experiments (see [32] for an example).

We should, however, point out that since the semiclassical states used in this context are not strict solutions of the constraints, the results obtained with them are only approximations of poorly controllable quality (see for example [15] for a discussion) and should not be interpreted as firm predictions of the theory. As initially stated, the situation is better in loop quantum cosmology, where semiclassical states that are physical, are available. As an example, the beautiful recent work [4] applies the ideas of quantum field theory on quantum space-time of [48, 49] in the context of loop quantum cosmology.

10.5 Outlook

Loop quantum gravity is a very unusual quantum field theory, and a promising approach to the unification of the principles of general relativity, and quantum theory. But open problems of great importance remain. We have in mind in particular the following questions:

- Are there restrictions on the types of matter that can be consistently coupled to gravity in the framework of loop quantum gravity?
- What role does the Barbero–Immirzi parameter ι play? Can its value be fixed by considerations other than black hole entropy?
- How can we extract physics from solutions of the Hamilton constraint?
- How can we obtain *controlled* approximations to the solutions of the dynamics?

Progress has already been made on all these. We think especially that the better understanding of the connection to spin foam models and the great results that have been achieved in loop quantum cosmology will help accelerate this progress in the near future.

Acknowledgments

The notes in this chapter are an extended version of my talk given at the workshop *Foundations of Space and Time – Reflections on Quantum Gravity* in honor of George Ellis at the STIAS in Stellenbosch (South Africa). I thank the organizers of that workshop for their wonderful hospitality, as well as for the great atmosphere they created during the workshop.

References

[1] I. Agullo, G. J. Fernando Barbero, E. F. Borja, J. Diaz-Polo, and E. J. S. Villasenor. The combinatorics of the SU(2) black hole entropy in loop quantum gravity. *Phys. Rev.*, **D80**:084006, 2009.

[2] A. Ashtekar. New variables for classical and quantum gravity. *Phys. Rev. Lett.*, **57**:2244–2247, 1986.

[3] A. Ashtekar, J. C. Baez, and K. Krasnov. Quantum geometry of isolated horizons and black hole entropy. *Adv. Theor. Math. Phys.*, **4**:1–94, 2000.

[4] A. Ashtekar, W. Kaminski, and J. Lewandowski. Quantum field theory on a cosmological, quantum space-time. *Phys. Rev.*, **D79**:064030, 2009.

[5] A. Ashtekar and J. Lewandowski. Differential geometry on the space of connections via graphs and projective limits. *J. Geom. Phys.*, **17**:191–230, 1995.

[6] A. Ashtekar and J. Lewandowski. Quantum theory of geometry. I: Area operators. *Class. Quant. Grav.*, **14**:A55–A82, 1997.

[7] A. Ashtekar and J. Lewandowski. Quantum theory of geometry. II: Volume operators. *Adv. Theor. Math. Phys.*, **1**:388–429, 1998.

[8] A. Ashtekar and Jerzy Lewandowski. Background independent quantum gravity: A status report. *Class. Quant. Grav.*, **21**:R53, 2004.

[9] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao, and T. Thiemann. Quantization of diffeomorphism invariant theories of connections with local degrees of freedom. *J. Math. Phys.*, **36**:6456–93, 1995.

[10] A. Ashtekar, T. Pawlowski, and P. Singh. Quantum nature of the big bang. *Phys. Rev. Lett.*, **96**:141301, 2006.

[11] G. J. Fernando Barbero. Real Ashtekar variables for Lorentzian signature space times. *Phys. Rev.*, **D51**:5507–10, 1995.

[12] M. Bojowald. Quantization ambiguities in isotropic quantum geometry. *Class. Quant. Grav.*, **19**:5113–230, 2002.

[13] M. Bojowald. Loop quantum cosmology. *Living Rev. Rel.*, **11**:4, 2008.

[14] M. Bojowald, G. Mortuza Hossain, M. Kagan, and S. Shankaranarayanan. Anomaly freedom in perturbative loop quantum gravity. *Phys. Rev.*, **D78**:063547, 2008.

[15] M. Bojowald, H. A. Morales-Tecotl, and H. Sahlmann. On loop quantum gravity phenomenology and the issue of Lorentz invariance. *Phys. Rev.*, **D71**:084012, 2005.

[16] J. Brunnemann and D. Rideout. Spectral analysis of the volume operator in loop quantum gravity. 2006.

[17] J. Brunnemann and D. Rideout. Properties of the volume operator in loop quantum gravity I: Results. *Class. Quant. Grav.*, **25**:065001, 2008.

[18] Y. Ding and C. Rovelli. The volume operator in covariant quantum gravity. 2009.

[19] B. Dittrich. Partial and complete observables for canonical general relativity. *Class. Quant. Grav.*, **23**:6155–84, 2006.

[20] B. Dittrich and J. Tambornino. A perturbative approach to Dirac observables and their space-time algebra. *Class. Quant. Grav.*, **24**:757–84, 2007.

[21] B. Dittrich and T. Thiemann. Testing the master constraint programme for loop quantum gravity. II: Finite dimensional systems. *Class. Quant. Grav.*, **23**:1067–88, 2006.

[22] B. Dittrich and T. Thiemann. Testing the master constraint programme for loop quantum gravity. IV: Free field theories. *Class. Quant. Grav.*, **23**:1121–42, 2006.

[23] B. Dittrich and T. Thiemann. Testing the master constraint programme for loop quantum gravity. V: Interacting field theories. *Class. Quant. Grav.*, **23**:1143–62, 2006.

[24] B. Dittrich and T. Thiemann. Are the spectra of geometrical operators in loop quantum gravity really discrete? *J. Math. Phys.*, **50**:012503, 2009.

[25] J. Engle, E. Livine, R. Pereira, and C. Rovelli. LQG vertex with finite Immirzi parameter. *Nucl. Phys.*, **B799**:136–49, 2008.

[26] J. Engle, R. Pereira, and C. Rovelli. The loop-quantum-gravity vertex-amplitude. *Phys. Rev. Lett.*, **99**:161301, 2007.

[27] J. Engle, A. Perez, and K. Noui. Black hole entropy and SU(2) Chern–Simons theory. 2009.

[28] C. Fleischhack. Representations of the Weyl algebra in quantum geometry. *Commun. Math. Phys.*, **285**:67–140, 2009.

[29] R. Gambini, J. Lewandowski, D. Marolf, and J. Pullin. On the consistency of the constraint algebra in spin network quantum gravity. *Int. J. Mod. Phys.*, **D7**:97–109, 1998.

[30] K. Giesel, S. Hofmann, T. Thiemann, and O. Winkler. Manifestly gauge-invariant general relativistic perturbation theory: I. Foundations. 2007.

[31] G. Immirzi. Real and complex connections for canonical gravity. *Class. Quant. Grav.*, **14**:L177–L181, 1997.

[32] F. R. Klinkhamer and M. Schreck. New two-sided bound on the isotropic Lorentz-violating parameter of modified-Maxwell theory. *Phys. Rev.*, **D78**:085026, 2008.

[33] K. Krasnov and C. Rovelli. Black holes in full quantum gravity. *Class. Quant. Grav.*, **26**:245009, 2009.

[34] J. Lewandowski and D. Marolf. Loop constraints: A habitat and their algebra. *Int. J. Mod. Phys.*, **D7**:299–330, 1998.

[35] J. Lewandowski, A. Okolow, H. Sahlmann, and T. Thiemann. Uniqueness of diffeomorphism invariant states on holonomy-flux algebras. *Commun. Math. Phys.*, **267**:703–33, 2006.

[36] A. Perez. On the regularization ambiguities in loop quantum gravity. *Phys. Rev.*, **D73**:044007, 2006.

[37] A. Perez and C. Rovelli. Physical effects of the Immirzi parameter. *Phys. Rev.*, **D73**:044013, 2006.

[38] M. P. Reisenberger and C. Rovelli. *Sum over surfaces* form of loop quantum gravity. *Phys. Rev.*, **D56**:3490–508, 1997.

[39] D. Jimenez Rezende and A. Perez. 4d Lorentzian Holst action with topological terms. *Phys. Rev.*, **D79**:064026, 2009.

[40] C. Rovelli. Area is the length of Ashtekar's triad field. *Phys. Rev.*, **D47**:1703–05, 1993.

[41] C. Rovelli. Comment on 'Are the spectra of geometrical operators in loop quantum gravity really discrete?' by B. Dittrich and T. Thiemann. 2007.

[42] C. Rovelli. Loop quantum gravity. *Living Rev. Rel.*, **11**:5, 2008.

[43] C. Rovelli and L. Smolin. Loop space representation of quantum general relativity. *Nucl. Phys.*, **B331**:80, 1990.
[44] C. Rovelli and L. Smolin. Discreteness of area and volume in quantum gravity. *Nucl. Phys.*, **B442**:593–622, 1995.
[45] C. Rovelli and L. Smolin. Spin networks and quantum gravity. *Phys. Rev.*, **D52**:5743–59, 1995.
[46] H. Sahlmann, T. Thiemann, and O. Winkler. Coherent states for canonical quantum general relativity and the infinite tensor product extension. *Nucl. Phys.*, **B606**:401–40, 2001.
[47] H. Sahlmann. Wave propagation on a random lattice. 2009.
[48] H. Sahlmann and T. Thiemann. Towards the QFT on curved spacetime limit of QGR. I: A general scheme. *Class. Quant. Grav.*, **23**:867–908, 2006.
[49] H. Sahlmann and T. Thiemann. Towards the QFT on curved spacetime limit of QGR. II: A concrete implementation. *Class. Quant. Grav.*, **23**:909–54, 2006.
[50] T. Thiemann. Kinematical Hilbert spaces for fermionic and Higgs quantum field theories. *Class. Quant. Grav.*, **15**:1487–512, 1998.
[51] T. Thiemann. QSD III: Quantum constraint algebra and physical scalar product in quantum general relativity. *Class. Quant. Grav.*, **15**:1207–47, 1998.
[52] T. Thiemann. QSD IV: 2+1 Euclidean quantum gravity as a model to test 3+1 Lorentzian quantum gravity. *Class. Quant. Grav.*, **15**:1249–80, 1998.
[53] T. Thiemann. QSD V: Quantum gravity as the natural regulator of matter quantum field theories. *Class. Quant. Grav.*, **15**:1281–314, 1998.
[54] T. Thiemann. Quantum spin dynamics (QSD). *Class. Quant. Grav.*, **15**:839–73, 1998.
[55] T. Thiemann. Quantum spin dynamics (QSD) II. *Class. Quant. Grav.*, **15**:875–905, 1998.
[56] T. Thiemann and O. Winkler. Gauge field theory coherent states (GCS). II: Peakedness properties. *Class. Quant. Grav.*, **18**:2561–636, 2001.
[57] T. Thiemann and O. Winkler. Gauge field theory coherent states (GCS) III: Ehrenfest theorems. *Class. Quant. Grav.*, **18**:4629–82, 2001.
[58] T. Thiemann. *Modern Canonical Quantum General Relativity*. Cambridge, UK: Cambridge University Press (2007) 819 pp.
[59] T. Thiemann. Gauge field theory coherent states (GCS). I: General properties. *Class. Quant. Grav.*, **18**:2025–64, 2001.
[60] T. Thiemann. The Phoenix project: Master constraint programme for loop quantum gravity. *Class. Quant. Grav.*, **23**:2211–48, 2006.
[61] R. M. Wald. The thermodynamics of black holes. *Living Rev. Rel.*, **4**:6, 2001.

11

Loop quantum gravity and cosmology

MARTIN BOJOWALD

> *"It would be permissible to look upon the Hamiltonian form as the fundamental one, and there would then be no fundamental four-dimensional symmetry in the theory. One would have a Hamiltonian built up from four weekly [sic] vanishing functions, given by [the Hamiltonian and diffeomorphism constraints]. The usual requirement of four-dimensional symmetry in physical laws would then get replaced by the requirement that the functions have weakly vanishing P.B.'s, so that they can be provided with arbitrary coefficients in the equations of motion, corresponding to an arbitrary motion of the surface on which the state is defined." P. A. M. Dirac, in "The theory of gravitation in Hamiltonian form," Proc. Roy. Soc. A* **246** *(1958) 333–43.*

11.1 Introduction

In its different incarnations, quantum gravity must face a diverse set of fascinating problems and difficulties, a set of issues best seen as both challenges and opportunities. One of the main problems in canonical approaches, for instance, is the issue of anomalies in the gauge algebra underlying space-time covariance. Classically, the gauge generators, given by constraints, have weakly vanishing Poisson brackets with each other: they vanish when the constraints are satisfied. After quantization, the same behavior must be realized for commutators of quantum constraints (or for Poisson brackets of effective constraints), or else the theory becomes inconsistent due to gauge anomalies. If and how canonical quantum gravity can be obtained in an anomaly-free way is an important question, not yet convincingly addressed in full generality. Posing one of the main obstacles to a complete formulation of

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

quantum gravity, this issue is hindering progress toward a detailed evaluation of quantum gravitational dynamics. A reliable phenomenological analysis must, after all, start with a consistent set of sufficiently general dynamical equations.

But the strong and tough requirement of anomaly-freedom is also an opportunity, for it allows an analysis of quantum space-time and the changes in its structure possibly implied by quantum gravity. Addressing the anomaly problem is, moreover, crucial for an understanding of the dynamics of quantum gravity, both in the sense of *constructing* consistent dynamical equations at the quantum level and in the sense of *analyzing* equations and their solutions to bring out physical effects.

Although the anomaly problem has not been addressed in full generality, several model systems have by now been analyzed in loop quantum cosmology, as reviewed in [30], with this question in mind. Loop quantum cosmology is a rather wide area within loop quantum gravity, analyzing several classes of model systems and perturbations around them. Loop quantum gravity, detailed in [5, 92, 105], is a canonical quantization of general relativity based on holonomies (the eponymous loops) as elementary variables. The use of holonomies allows a background-independent formulation free of auxiliary metrics, and it implies several specific properties of the resulting dynamics.

In all the systems used in loop quantum cosmology, quantization techniques close to those of a general loop quantization are used; they can thus be seen as capturing at least some of the crucial properties of full loop quantum gravity. To different degrees, most of these models make additional use of symmetry reduction as introduced in [36], simplifying much of the quantum geometry and thereby providing rather direct access to the much less understood quantum dynamics. Thanks to these steps, implications of the quantum dynamics of loop quantum gravity, as generally defined based on [94, 103], have been evaluated quite explicitly for the first time.

In general terms as well as for particular questions arising in loop quantum cosmology and loop quantum gravity, three key issues regarding quantum space-time arise.

Effective dynamics: In quantum gravity, geometry is described unsharply by whole states with all their fluctuations and correlations in addition to the expectation values for an average geometry. Equations of motion for expectation values receive quantum corrections in their effective dynamics, as it may describe an effective geometry. Along with this come not only new mathematical space-time structures but also a vast enlargement of the number of degrees of freedom by quantum variables.

The highest control of such a high-dimensional dynamics is usually obtained for dynamical coherent states, defined as states saturating uncertainty relations at all times. If such states exist, they provide insights into the minimal deviations from classical behavior expected for a quantum system. The form and behavior of dynamical coherent states in loop quantum cosmology can be highlighted in several models, bringing out the role of space-time fluctuations and correlations as degrees of freedom beyond the classical ones. Exact dynamical coherent states exist only in special models and for specific initial values. Nevertheless, they allow interesting views on the generic quantum dynamics as it arises in quantum gravity.

Before all corrections are derived for a large class of models, a clear analysis provided by dynamical coherent states, when they exist, unambiguously shows the first deviations from classical behavior. More generally, when exact dynamical coherent states do not exist, additional quantum corrections will result. They can often be computed perturbatively, analogously to loop corrections in interacting quantum field theories.

Discrete dynamics: In addition to those generic effects due to non-classical state parameters, underlying space-time structures typically change even for the expectation values of quantum gravity states. Most importantly, discrete geometry, at least in purely spatial terms which is by now well understood in loop quantum gravity following [3, 4, 95], shows several detailed properties of importance for the dynamics and thus for space-time geometry.

A spatial slice in space-time is equipped with a discrete quantum geometry, roughly seen as making space built from atomic patches of certain discrete sizes. One of the main problems of quantum dynamics is to show how these spatial atoms along slices fit together to form a quantum space-time – or, more dynamically, how the spatial atoms change, merge, subdivide and interact as time is let loose. For an expanding universe, one would expect the discrete spatial structure to be refined as the volume increases; otherwise discrete sizes would be enlarged by huge factors, especially during inflation, making them macroscopic. The full dynamics of loop quantum gravity has indeed provided several hints that the number of discrete building blocks must change from slice to slice, once the dynamics is implemented consistently. This lattice refinement can be modeled even in the simplest, most highly symmetric situations of loop quantum cosmology, laid out in [24, 28]. And it has shown several specific implications by which its precise form can already be constrained.

Consistent dynamics: Dynamics unfolds in time, but time is relative. Making sure that descriptions using different notions of time, corresponding to measurements by different observers, can agree about their physical insights requires the consistency conditions of general covariance. Since general covariance in gravity is implemented as gauge transformations, any quantization or even just a modification of the theory must respect this and be anomaly-free.

While the previous two points manifest themselves already in homogeneous models, where they can most easily be studied, the anomaly problem arises only in inhomogeneous situations. Within homogeneity, all but one of the gauge transformations underlying covariance are fixed. Anomalies can only arise if there are at least two independent gauge transformations; after quantization, they would be anomalous if their composition is no longer a gauge transformation. One could, of course, make it a gauge transformation by definition, by declaring the whole group generated by the independent gauge transformations as the gauge group. But if this group is too large, it would identify variables which are to be considered physically distinct, removing observables and degrees of freedom and in many cases leaving no non-trivial solutions. Changing the gauge transformations of a classical theory by quantum effects requires much care; only so-called consistent deformations of the classical gauge generators can provide well-defined quantizations.

Addressing these questions is crucial, not only for a complete formulation of quantum gravity but also for reliable cosmological applications based on the resulting set of equations (such as singularity removal or structure formation). In what follows, we will describe the current status based on the models of loop quantum cosmology.

11.2 Effective dynamics

In a general sense, effective equations of a quantum system describe the behavior of expectation values in a state. Deriving such equations for the expectation values of basic operators, such as $\hat{q}$ and $\hat{p}$ in quantum mechanics, shows how quantum effects change the classical equations of motion. If the equations can be solved or analyzed, the manifestation of quantum behavior will be seen. Such effects play an important role in any interacting quantum theory, so also in quantum gravity and especially in quantum cosmology which devotes itself to the analysis of extremely long evolution times. During those times, quantum states may change drastically, and quantum corrections grow to be significant.

11.2.1 Momentous quantum mechanics

Effective equations must be state-dependent since they are equations for expectation values in a state with many more independent parameters. In general, quantum fluctuations will influence the behavior of expectation values, and must then be included in effective equations in some form. More generally, and following [38], we can parameterize a state by all its moments

$$G^{a,b} \equiv G^{\overbrace{q\cdots q}^{a}\overbrace{p\cdots p}^{b}} = \langle (\hat{q} - \langle \hat{q} \rangle)^a (\hat{p} - \langle \hat{p} \rangle)^b \rangle_{\mathrm{Weyl}} \tag{11.1}$$

defined for $a + b \geq 2$, in addition to the expectation values $\langle \hat{q} \rangle$ and $\langle \hat{p} \rangle$. (The subscript "Weyl" denotes totally symmetric ordering. We will use the two notations indicated on the left interchangeably, at least for small $a + b$.) For instance, $G^{2,0} \equiv G^{qq}$ is the square of position fluctuations, and $G^{1,1} \equiv G^{qp} = \frac{1}{2}\langle \hat{q}\hat{p} + \hat{p}\hat{q} \rangle - \langle \hat{q} \rangle \langle \hat{p} \rangle$ the covariance. Moments are not arbitrary, most importantly being restricted by uncertainty relations such as

$$G^{qq} G^{pp} - (G^{qp})^2 \geq \frac{\hbar^2}{4} \,. \tag{11.2}$$

For pure states, the set of moments as defined here is overcomplete; the framework more generally allows for mixed states, too.

All the moments are dynamical. Given a Hamiltonian operator, for every observable $\hat{O}$ we can derive an equation of motion

$$\frac{\mathrm{d}\langle \hat{O} \rangle}{\mathrm{d}t} = \frac{\langle [\hat{O}, \hat{H}] \rangle}{i\hbar} \tag{11.3}$$

of its expectation value. Specific examples are obtained for the terms in a moment, and so we can derive their equations of motion. For the square of position fluctuations, for instance, we have

$$\frac{\mathrm{d}G^{qq}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\langle \hat{q}^2 \rangle - \langle \hat{q} \rangle^2) = \frac{\langle [\hat{q}^2, \hat{H}] \rangle}{i\hbar} - 2\langle \hat{q} \rangle \frac{\langle [\hat{q}, \hat{H}] \rangle}{i\hbar} .$$

Introducing Poisson brackets on the space of moments via

$$\{\langle \hat{A} \rangle, \langle \hat{B} \rangle\} = \frac{\langle [\hat{A}, \hat{B}] \rangle}{i\hbar} , \tag{11.4}$$

extended by linearity and the Leibniz rule, equations of motion take Hamiltonian form:

$$\frac{\mathrm{d}G^{a,b}}{\mathrm{d}t} = \{G^{a,b}, H_Q\} \tag{11.5}$$

with the quantum Hamiltonian $H_Q := \langle \hat{H} \rangle$.

In general, the moments are all coupled to each other and to the equations for expectation values. One can see this by expanding

$$
\begin{aligned}
H_Q(\langle\hat{q}\rangle, \langle\hat{p}\rangle, G^{a,b}) &= \langle H(\hat{q}, \hat{p})\rangle = \langle H(\langle\hat{q}\rangle + (\hat{q} - \langle\hat{q}\rangle), \langle\hat{p}\rangle + (\hat{p} - \langle\hat{p}\rangle))\rangle \\
&= H(\langle\hat{q}\rangle, \langle\hat{p}\rangle) + \sum_{a,b:a+b\geq 2} \frac{1}{a!b!} \frac{\partial^{a+b} H(\langle\hat{q}\rangle, \langle\hat{p}\rangle)}{\partial\langle\hat{q}\rangle^a \partial\langle\hat{p}\rangle^b} G^{a,b}
\end{aligned}
\tag{11.6}
$$

(where we assumed the Hamiltonian operator $\hat{H}$ to be Weyl-ordered in $\hat{q}$ and $\hat{p}$) and noticing the coupling terms of expectation values and moments for any non-quadratic potential. The Hamiltonian flow (11.3) or (11.5) then couples expectation values $\langle\hat{q}\rangle$ and $\langle\hat{p}\rangle$ to all the moments.

At this stage, we have an exact but usually horribly complicated Hamiltonian description of quantum evolution. The partial differential equation for a state in Schrödinger's formulation is replaced by infinitely many ordinary differential equations for the moments. Most of the labor that goes into deriving treatable effective equations consists in extracting the required information about expectation values without having to solve for a full state, or all its moments, and to specify the regimes where this is reliable. In a semiclassical approximation based on near-Gaussian states, for instance, a moment of order $a + b$ is typically of the order $\hbar^{(a+b)/2}$, giving rise to a natural expansion in powers of $\hbar$. To any given order, only finitely many moments need be considered.[1]

11.2.2 Harmonic oscillator

In special systems, equations of motion for the moments decouple automatically to finite sets. The best known case is the harmonic oscillator, whose quantum Hamiltonian is

$$
H_Q = \frac{1}{2m}\langle\hat{p}\rangle^2 + \frac{1}{2}m\omega^2\langle\hat{q}\rangle^2 + \frac{1}{2}m\omega^2 G^{qq} + \frac{1}{2m}G^{pp}. \tag{11.7}
$$

Second-order moments appear, but they do not couple to the expectation values. They rather provide the zero-point energy due to quantum fluctuations. Hamiltonian

[1] Truncating the phase space in this way leads to Poisson manifolds spanned by the moments to a certain order. In general, these Poisson structures are degenerate; for instance, there are three independent second-order moments, forming a space which cannot carry a non-degenerate Poisson structure. Effective equations thus make crucial use of Poisson geometry, not symplectic geometry as in other geometric formulations of quantum mechanics such as the one going back to [77].

equations of motion are only finitely coupled,

$$\begin{aligned}
\frac{\mathrm{d}\langle\hat{q}\rangle}{\mathrm{d}t} &= \{\langle\hat{q}\rangle, H_Q\} = \frac{1}{m}\langle\hat{p}\rangle \\
\frac{\mathrm{d}\langle\hat{p}\rangle}{\mathrm{d}t} &= \{\langle\hat{p}\rangle, H_Q\} = -m\omega^2\langle\hat{q}\rangle \\
\frac{\mathrm{d}G^{a,b}}{\mathrm{d}t} &= \{G^{a,b}, H_Q\} = \frac{1}{m}aG^{a-1,b+1} - m\omega^2 bG^{a+1,b-1}.
\end{aligned} \tag{11.8}$$

For the expectation values we have exactly the classical equations without quantum corrections, as is well known for the harmonic oscillator. Here, effective equations for $\langle\hat{q}\rangle$ and $\langle\hat{p}\rangle$ are identical to the classical ones.

The remaining equations then show how the state evolves once initial values for moments have been chosen. For stationary states, for instance, a vanishing covariance G^{qp} ensures that fluctuations are time-independent. The covariance is constant in time if $\dot{G}^{qp} = G^{pp}/m - m\omega^2 G^{qq} = 0$, or $G^{pp} = m^2\omega^2 G^{qq}$. With this, the non-classical contribution to the energy H_Q in (11.7) is $m\omega^2 G^{qq}$, where $G^{qq} \geq \hbar/2m\omega$ from (11.2). A minimal two-point energy $\frac{1}{2}\hbar\omega$ results for the ground state, derived purely by effective means (even though the ground state is not at all semiclassical).

Decoupled equations of this form are very useful because they can directly show important aspects of quantum dynamics. Expectation value equations decouple from the rest whenever the algebra of basic operators together with the Hamiltonian is linear: in this case, the time derivative of the expectation value of any one of the basic operators is an expectation value of a basic operator. Such systems are solvable in a strong sense; there is no quantum back-reaction from the moments on the dynamics of expectation values. And the dynamics of moments of a given order depends only on moments of the same order. In terms of quantum field theory, solvable models correspond to free theories.

11.2.3 Low-energy effective potential

With interactions, or for anharmonic terms, quantum back-reaction results. Moments couple non-trivially to expectation values and become important for their dynamics. In this way, the state dependence of effective equations ensues. The system of equations necessarily becomes of higher dimension than classically, with new dynamical quantum degrees of freedom given by the moments.

For a quantum mechanical system with an anharmonic potential, for instance, we have the classical Hamiltonian $H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 + U(q)$ with anharmonicity

$U(q)$. In terms of dimensionless quantum variables

$$g^{a,b} = \hbar^{-(a+b)/2}(m\omega)^{a/2-b/2}G^{a,b} \tag{11.9}$$

the quantum Hamiltonian is

$$\begin{aligned} H_Q = &\frac{1}{2m}\langle\hat{p}\rangle^2 + \frac{1}{2}m\omega^2\langle\hat{q}\rangle^2 + U(\langle\hat{q}\rangle) + \frac{\hbar\omega}{2}(g^{qq} + g^{pp}) \\ &+ \sum_{n=2}^{\infty}\frac{1}{n!}\left(\frac{\hbar}{m\omega}\right)^{n/2} U^{(n)}(\langle\hat{q}\rangle)g^{q^n} \end{aligned} \tag{11.10}$$

and generates equations of motion

$$\begin{aligned} \frac{\mathrm{d}\langle\hat{q}\rangle}{\mathrm{d}t} &= \frac{1}{m}\langle\hat{p}\rangle \\ \frac{\mathrm{d}\langle\hat{p}\rangle}{\mathrm{d}t} &= -m\omega^2\langle\hat{q}\rangle - U'(\langle\hat{q}\rangle) - \sum_{n=2}^{\infty}\frac{1}{n!}\left(\frac{\hbar}{m\omega}\right)^{n/2} U^{(n+1)}(\langle\hat{q}\rangle)g^{q^n} \\ \dot{g}^{a,b} &= a\omega g^{a-1,b+1} - b\omega g^{a+1,b-1} - \frac{bU''(\langle\hat{q}\rangle)}{m\omega}g^{a+1,b-1} \\ &\quad + \frac{b\sqrt{\hbar}U'''(\langle\hat{q}\rangle)}{2(m\omega)^{3/2}}\left(g^{a,b-1}g^{qq} - g^{a+2,b-1} + \frac{(b-1)(b-2)}{12}g^{a,b-3}\right) \\ &\quad + \frac{b\hbar U''''(\langle\hat{q}\rangle)}{6(m\omega)^2}\left(g^{a,b-1}g^{qqq} - g^{a+3,b-1} + \frac{(b-1)(b-2)}{4}g^{a+1,b-3}\right) + \cdots \end{aligned} \tag{11.11}$$

The leading quantum correction appears in the equation for $\langle\hat{p}\rangle$ at order $\hbar$ ($n = 2$), for which we have to know the position fluctuation g^{qq}. In general, this is an independent variable subject to its own equation of motion. Its evolution couples to other quantum variables, eventually coupling the whole infinite system. At this stage, approximations are required. For semiclassical states, we may drop terms of higher order in $\hbar$,[2] providing a closed system of effective equations

$$\begin{aligned} \frac{\mathrm{d}\langle\hat{q}\rangle}{\mathrm{d}t} &= \frac{1}{m}\langle\hat{p}\rangle \\ \frac{\mathrm{d}\langle\hat{p}\rangle}{\mathrm{d}t} &= -m\omega^2\langle\hat{q}\rangle - U'(\langle\hat{q}\rangle) - \frac{1}{2}\frac{\hbar}{m\omega}U'''(\langle\hat{q}\rangle)g^{qq} + O(\hbar^{3/2}) \end{aligned}$$

[2] By definition (11.9), the leading semiclassical $\hbar$-dependence of $g^{a,b}$ has been factored out, such that explicit factors of $\hbar$ in the equations of motion suffice to read off orders. In terms of the original moments, at second order we ignore all terms $\hbar^n G^{a,b}$ with $2n + a + b > 2$.

$$\frac{\mathrm{d}g^{qq}}{\mathrm{d}t} = 2\omega g^{qp}$$
$$\frac{\mathrm{d}g^{qp}}{\mathrm{d}t} = \omega(g^{pp} - g^{qq}) - \frac{U''(\langle\hat{q}\rangle)}{m\omega} g^{qq} + O(\sqrt{\hbar})$$
$$\frac{\mathrm{d}g^{pp}}{\mathrm{d}t} = -2\omega g^{qp} - 2\frac{U''(\langle\hat{q}\rangle)}{m\omega} g^{qp} + O(\sqrt{\hbar})$$

for expectation values and second-order moments. With higher moments dropped, this system shows the leading quantum corrections in a finitely coupled system.

For anharmonic oscillators, a further treatment is possible which makes use of an adiabatic approximation: we assume that time derivatives of the moments are small compared to the other terms. In this case, equations of motion for moments become algebraic relationships between them. For the leading adiabatic order $g_0^{a,b}$, combined with the previous $\hbar$-expansion, we must solve

$$0 = \{g_0^{a,b}, H_Q\} = \omega\left(a g_0^{a-1,b+1} - b\left(1 + \frac{U''(\langle\hat{q}\rangle)}{m\omega^2}\right) g_0^{a+1,b-1}\right) + O(\sqrt{\hbar})$$

which is of the stationary harmonic form, but with position-dependent coefficients. Although $\mathrm{d}\langle\hat{q}\rangle/\mathrm{d}t \neq 0$, this is consistent with the adiabatic assumption. The general solution is

$$g_0^{a,b} = \binom{(a+b)/2}{b/2}\binom{a+b}{b}^{-1}\left(1 + \frac{U''(\langle\hat{q}\rangle)}{m\omega^2}\right)^{b/2} g_0^{a+b,0}$$

for even a and b, and $g_0^{a,b} = 0$ whenever a or b are odd. The values of $g_0^{n,0}$ for n even remain free, but must satisfy a condition following from the next adiabatic order $g_1^{a,b}$.

For the first adiabatic order, we now consider $g_0^{a,b}$ weakly time-dependent via $\langle\hat{q}\rangle$, but assume $g^{a,b} - g_0^{a,b}$ time-independent. Solutions to the resulting equations $\{g_1^{a,b} - g_0^{a,b}, H_Q\} = 0$ provide the first adiabatic order[3] $g_1^{a,b}$, whose time dependence is given by derivatives of the zeroth (adiabatic) order moments:

$$\{g_1^{a,b}, H_Q\} = \omega\left(a g_1^{a-1,b+1} - b\left(1 + \frac{U''(\langle\hat{q}\rangle)}{m\omega^2}\right) g_1^{a+1,b-1}\right) = \dot{g}_0^{a,b}.$$

[3] The nth adiabatic order $g_n^{a,b}$ is assumed to satisfy $\{g_n^{a,b} - g_{n-1}^{a,b}, H_Q\} = 0$ and thus solves $\{g_n^{a,b}, H_Q\} = \dot{g}_{n-1}^{a,b}$. Iterating over n, this provides algebraic equations for all orders.

This equation implies

$$\sum_{b\,\mathrm{even}} \binom{(a+b)/2}{b/2} \left(1+\frac{U''(\langle\hat{q}\rangle)}{m\omega^2}\right)^{\frac{b}{2}} \dot{g}_0^{a,b} = 0$$

and requires $g_0^{n,0} = C_n(1+U''(\langle\hat{q}\rangle)/m\omega^2)^{-n/4}$. The remaining constant C_n can be fixed by requiring the moments to be of harmonic oscillator form for $U=0$: $C_n = 2^{-n}n!/(n/2)!$. In particular, $g_0^{qq} = \frac{1}{2}(1+U''(\langle\hat{q}\rangle)/m\omega^2)^{-1/2}$, and the quantum correction to the effective force is $-\frac{1}{4}(\hbar/m\omega)U'''(\langle\hat{q}\rangle)(1+U''(\langle\hat{q}\rangle)/m\omega^2)^{-1/2}$ as it arises from an effective potential

$$V_{\mathrm{eff}}(q) = \frac{1}{2}m\omega^2 q^2 + U(q) + \frac{1}{2}\hbar\omega\sqrt{1+\frac{U''(q)}{m\omega^2}} \qquad (11.12)$$

from $\frac{1}{2}\hbar\omega(g_0^{qq}+g_0^{pp}) + \frac{1}{2}(\hbar/m\omega)U''(\langle\hat{q}\rangle)g_0^{qq}$ in (11.10). This function agrees with path integral calculations of the low-energy effective action as derived by [54]. Here, only the zeroth adiabatic order combined with first order in $\hbar$ has been used. To second order computed by [38], there is also a correction for the mass term, still in agreement with [54].

The adiabatic approximation is an example for regimes where the quantum evolution of moments can be further reduced to result in explicit effective forces. While the forces come from coupling terms between expectation values and quantum variables such as fluctuations, the latter do not appear explicitly. Quantum effects then manifest themselves indirectly in the form of effective terms depending on the expectation values, their quantum origin being indicated only by the presence of $\hbar$ but not by explicit quantum degrees of freedom. Such effective terms cannot always be derived since some regimes, where effective descriptions may well apply, do require the larger freedom of higher-dimensional effective systems coupling quantum degrees of freedom explicitly. For low-energy effective potentials, the adiabatic approximation is responsible for the reduction to a system of classical form as far as degrees of freedom are concerned.

From the derivation we can also see how the state dependence is a crucial part of effective equations, even though the final expression (11.12) for the potential seems state-independent. To fix all free constants in effective equations for anharmonic systems, in particular C_2, we had to refer to the harmonic oscillator vacuum. We are thus expanding around the vacuum of the free, solvable theory whose moments are known. Quantum corrections from the interacting vacuum have been derived in passing, by obtaining the leading adiabatic orders of g^{qq}. For the low-energy effective potential, this result was reinserted in the equations of motion, somewhat hiding the state dependence.

With the higher-dimensional viewpoint of effective systems, keeping some of the moments as independent parameters, we obtain extra information about the interacting theory not seen in a simple effective potential. For instance, from our intermediate calculations we directly have $g_0^{qq} = \frac{1}{2}(1 + U''(\langle\hat{q}\rangle)/m\omega^2)^{-1/2}$, while $g_0^{pp} = (1 + U''(\langle\hat{q}\rangle)/m\omega^2)g_0^{qq} = \frac{1}{2}(1 + U''(\langle\hat{q}\rangle)/m\omega^2)^{1/2}$ and $g_0^{qp} = 0$. To zeroth adiabatic order, the interacting ground state keeps saturating the uncertainty relation, but fluctuations are no longer exactly constant.

11.2.4 Quantum cosmology

In relativistic systems, there is no absolute time with evolution generated by a Hamiltonian. Rather, relativistic systems are subject to a Hamiltonian constraint C. It generates arbitrary changes of the time coordinate as gauge transformations $\delta_\varepsilon f = \varepsilon\{f, C\}$ for phase space functions f. From observable quantities O left unchanged by gauge transformations, i.e. $\{O, C\} = 0$, dynamical properties then follow. Since the invariance condition $\{O, C\} = 0$ removes one dimension from the initial phase space, for consistency we must require $C = 0$ as a constraint.[4]

Constrained formulations can be introduced also for non-relativistic systems by parameterization, adding a time degree of freedom t with momentum p_t and replacing the Hamiltonian H by the Hamiltonian constraint $C = p_t - H$. For time-independent Hamiltonians, p_t is gauge-invariant while the gauge transformation of t is $\delta_\varepsilon t = \varepsilon\{t, C\} = \varepsilon$; t can be changed at will. For the remaining observables,

$$0 = \{O, C\} = \frac{\mathrm{d}O}{\mathrm{d}t} - \{O, H\}$$

imposes Hamilton's equations of motion.

For relativistic quantum systems, the effective techniques described so far cannot directly be applied. There are extensions to effective constraints briefly described later, as developed in [42, 48]. But more simply, effective equation techniques can be used if systems are first deparameterized, reverting the above procedure. If there is a variable ϕ, then called an internal time variable, such that the Hamiltonian constraint can be written as $C = p_\phi - H(q, p)$ with the momentum p_ϕ of ϕ and

[4] Factoring out the Hamiltonian flow generated by the constraint C via its Hamiltonian vector field $X_C = \{\cdot, C\}$, we obtain a projection $\pi : M \to M/X_C$ from the original phase space M by identifying all points along the orbits of X_C. All observables O naturally descend to the factor space since they are constant along the orbits, and so does C. In this way, we obtain a complete set of functions on M/X_C. On the factor space, we have a natural Poisson structure $\{f, g\}_{M/X_C} = \{\pi^* f, \pi^* g\}_M$, pushing forward the Poisson bivector via π. This Poisson structure is degenerate: $\{O, C\}_{M/X_C} = 0$ for all functions O on M/X_C. The constraint C becomes a Casimir function on the factor space, and symplectic leaves of the Poisson structure are given by $C = \text{const}$. Any leaf carries a non-degenerate symplectic structure and can be taken as a reduced phase space, but $C = 0$ is distinguished. In this case, we can write gauge transformations as $\delta_\varepsilon f = \{f, \varepsilon C\}$, even for phase-space functions ε. More details of Poisson geometry in the context of constrained systems are described by [40].

a function H independent of ϕ and p_ϕ, gauge transformations generated by the constraint take the form

$$\delta_\varepsilon f = \varepsilon\{f(q,p,\phi,p_\phi),C\} = \varepsilon\left(\frac{\partial f}{\partial\phi} - \{f,H\}\right).$$

Gauge-invariant quantities of the theory are thus those evolving in the usual Hamiltonian way as generated by the ϕ-Hamiltonian H.

Once deparameterized, the observables of a constrained system can be derived by analyzing an ordinary Hamiltonian flow. At this stage, effective techniques as described before can be applied to quantizations of deparameterized models. Effective equations of motion derived from

$$\frac{\mathrm{d}\langle\hat{O}\rangle}{\mathrm{d}\phi} = \frac{\langle[\hat{O},\hat{H}]\rangle}{i\hbar} \tag{11.13}$$

then provide means to solve for quantum observables $\langle\hat{O}\rangle(\phi)$ and their physical evolution. For the initial constrained system, solutions $\langle\hat{O}\rangle(\phi)$ are observables as functions on the full phase space including (ϕ, p_ϕ). (Remaining phase space variables enter the expression $\langle\hat{O}\rangle(\phi)$ via initial values taken for them when solving the differential equations (11.13).) The deperameterization endows $\langle\hat{O}\rangle(\phi)$ with the interpretation as an observable $\langle\hat{O}\rangle$ evolving with respect to the internal time ϕ. This is the relational picture for interpreting constrained dynamics, developed classically in [17, 60, 61, 90, 91].

Examples for such systems in cosmology are homogeneous models sourced by a free, massless scalar. Its energy density is purely kinetic, $\rho = \frac{1}{2}a^{-6}p_\phi^2$, such that the Friedmann equation, solved for p_ϕ, provides a Hamiltonian for ϕ-evolution. For a spatially flat Friedmann–Robertson–Walker model, the ϕ-Hamiltonian then turns out to be quadratic in suitable canonical variables: for instance, the Hubble parameter $\mathcal{H} = \dot{a}/a$ is canonically conjugate to the volume $V = a^3$, and the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho = \frac{4\pi G}{3}\frac{p_\phi^2}{a^6}$$

tells us that $p_\phi^2 \propto V^2\mathcal{H}^2$. Upon taking a square root, we have a quadratic ϕ-Hamiltonian. Based on this observation, state properties have been determined by [25, 26]. (Strictly speaking, the ϕ-Hamiltonian is of the form $|qp|$ which is not quadratic. However, for effective equations one can show that the absolute value can be dropped, providing a linear quantum system. We only have to require an initial state to be supported on a definite part of the spectrum of $\widehat{qp}$, either the positive or the negative one, which is then preserved in time since $\hat{H}$ is preserved. The absolute value then just amounts to multiplication with ± 1. For initially semiclassical states, this requirement does not lead to restrictions for low-order moments.)

For other systems, perturbation theory can be used as described above for anharmonic oscillators. A crucial difference, according to the analysis in [44], is that no adiabatic regime has so far been found for quantum cosmology, blocking the complete expression of quantum variables in terms of an effective potential. On the other hand, higher-dimensional effective systems, where quantum variables are taken as independent variables subject to their own evolution, can be analyzed and show how states back-react on the expectation value trajectories.

If we include a mass term or a potential for the scalar, the system becomes "time-dependent" in ϕ. Extra care is required, but perturbation theory still applies for small and flat potentials. The justification for this procedure in the time-dependent case comes from an extension of the effective equation procedure to constrained systems introduced by [42, 48] without requiring deparameterization. Quantum constraint operators then imply the existence of infinitely many constraints $\langle \widehat{\mathrm{pol}} \hat{C} \rangle$ on the quantum phase space, in general all independent for different polynomials $\widehat{\mathrm{pol}}$ in the basic operators. This large number of constraints restricts not only expectation values (as the classical analog) but also the corresponding quantum variables. A complete reduction to the physical state space results, without requiring a deparameterization to exist. With these general techniques, effective equations for relativistic systems are thus fully justified. As discussed in Section 11.4, inhomogeneous models bring in a new level due to the anomaly problem, which does not arise for a single classical constraint.

11.2.5 Symplectic structure

Quantum corrections in canonical effective equations come from corrected Hamiltonians or corrected constraints as expectation values of Hamiltonian (constraint) operators. The Hamiltonian H or constraints C_I, together with Lagrange multipliers N^I (for gravity lapse and shift multiplying the Hamiltonian and diffeomorphism constraints), form an important contribution to the action. But the action

$$S[q(t), p(t); N^I] = \int \mathrm{d}t \left(\dot{q} p - H(q, p) - N^I C_I(q, p) \right)$$

of a Hamiltonian system in canonical form has an extra contribution, the one that determines the symplectic structure between configuration and momentum variables. One might wonder whether it is enough to look for quantum corrections in the Hamiltonian or the constraints without correcting the symplectic structure.

With symplectic structure corrections, an effective action might take a different form than suggested by an analysis only of Hamiltonians and constraints.

There might indeed be corrections to the symplectic structure, but they would follow from the same algebraic notions used for canonical effective equations. Poisson brackets of quantum variables are given by expectation values of commutators as used before; see (11.4). Any potential corrections to the symplectic structure have thus been taken care of by applying that formula consistently. If there are no changes in Poisson bracket relations for effective equations, it is only because they do not change for expectation values of basic operators: they have commutator relations mimicking the classical Poisson algebra, which is linear for the basic objects. Taking expectation values of these linear structures does not lead to corrections. For canonical basic operators $\hat{q}$ and $\hat{p}$, for instance, we have $\{\langle\hat{q}\rangle, \langle\hat{p}\rangle\} = \langle[\hat{q}, \hat{p}]\rangle / i\hbar = 1$.

The Poisson structure changes only because the dimension of the phase space increases by the quantum variables $G^{a,b}$. These variables satisfy Poisson relationships following from the quantum theory, but this does not affect the Poisson brackets of expectation values of basic operators. In an action, the symplectic term for classical variables remains unchanged. But if a higher-dimensional effective system with independent quantum variables is used, their symplectic terms add to the action. The symplectic structure is only extended to include new degrees of freedom; it is not quantum corrected.

Sometimes, one can make assumptions[5] about the dependence of quantum variables on expectation values, or even derive those by an adiabatic approximation. If this is done, one can insert such expressions, schematically $G^I(q,p)$, into the symplectic form for quantum variables, of the form $\Omega_{IJ}(G)\mathrm{d}G^I \wedge \mathrm{d}G^J$ where $\Omega_{IJ}(G)$ follows from (11.4) purely kinematically. Then, a term $2\Omega_{IJ}(\partial G^I/\partial q)(\partial G^J/\partial p)\mathrm{d}q \wedge \mathrm{d}p$ results which would add to the symplectic term of expectation values; $\dot{q}p$ becomes $(1+2\Omega_{IJ}(\partial G^I/\partial q)(\partial G^J/\partial p))\dot{q}p$. Now, the correction has dynamical information via the partial solutions $G^I(q,p)$. (For the specific example of anharmonic oscillators, no such corrections arise since the effective equations for g^{ab}, and thus their solutions, depend only on $\langle\hat{q}\rangle$, not $\langle\hat{p}\rangle$.) With independent quantum variables, as they are most often required, no symplectic structure corrections result. Expectation values of Hamiltonians or constraints are then the key source for quantum corrections.

For quantum gravity, we expect higher curvature terms in an effective action, and thus higher derivative terms. By the preceding discussion, this can only come from independent moments $G^{a,b}$. But quantum gravity also has other implications,

[5] A quantum cosmological model with assumptions for semiclassical states has been analyzed in [102].

such as the emergence of discrete spatial structures expected in particular from loop quantum gravity. They, too, must affect the effective dynamics.

11.3 Discrete dynamics

Our main interest from now on will be to apply effective equation techniques to canonical quantum gravity. A strict derivation requires detailed knowledge of the mathematical properties of operators involved, as well as information about the representation of states. Effective equations and effective constraints are, after all, obtained via expectation values of Hamiltonians or Hamiltonian constraint operators. These operators are first constructed within the full quantum theory of gravity or one of its models, and the resulting constructs will be sensitive to properties of the basic operators, analogous to $\hat{q}$ and $\hat{p}$ in quantum mechanics. Some of these properties will descend to the effective level once expectation values of Hamiltonians are computed, and then affect also the effective dynamics. At this stage, general effective techniques naturally tie in with specific constructions of a concrete quantum theory at hand.

11.3.1 Loop quantum gravity

Classically, the canonical structure of general relativity is given by the spatial part q_{ab} of the space-time metric g_{ab}, as well as momenta related to its change in time, or the extrinsic curvature K_{ab} of the spatial slices Σ. These specific quantities are meaningful only when a choice for time, a time function t such that $\Sigma: t =$ const, has been made, and so this formalism is often seen as breaking covariance. But space-time covariance is broken only superficially and is restored when all the dynamical constraints have been solved – in addition to a local Hamiltonian constraint, three components of the diffeomorphism constraint to generate all four independent space-time coordinate changes. The theory is, after all, equivalent to general relativity in its Lagrangian form; just formulating it in different variables cannot destroy underlying symmetries. In fact, the Hamiltonian formulation still shows the full generators of all gauge transformations in explicit form, by the constraints it implies for the fields. By the general theory of constraint analysis, a discussion of gauge at the Hamiltonian level is then even more powerful than at the Lagrangian one.

Smearing

Before addressing quantum constraints, the quantum theory must be set up. For a well-defined quantization one turns the basic fields into operators such that the classical Poisson bracket is reflected in commutator relationships. In constructions of quantum field theories one has to face the problem that Poisson brackets of the fields involve delta functions since they are non-vanishing only when the values of two conjugate fields are taken at the same point, as in $\{\phi(x), p_\phi(y)\} = \delta(x, y)$ for a scalar field. A simple but powerful remedy is to "smear" the fields by integrating them against test functions over space. Again for a scalar field, we could use $\phi[\mu] := \int d^3x \sqrt{\det q}\, \mu(x)\phi(x)$ for which $\{\phi[\mu], p_\phi(y)\} = \int d^3x \mu(x)\delta(x, y) = \mu(y)$. The Poisson algebra for the smeared fields is free of delta functions thanks to the integration. Moreover, for sufficiently general classes of test functions $\mu(x)$, smeared fields capture the full information contained in the local fields and can be used to set up a general theory.

After smearing, a well-defined Poisson algebra of basic objects results, ready to be turned into an operator algebra by investigating its representations. For gravity, we would use smeared versions of the spatial metric and its change in time. But here, a second problem arises. One of the dynamical fields to be quantized is the spatial metric, but to smear fields we need a metric for the integration measure. This is no problem when fields to be quantized are non-gravitational. As with the scalar in the example above, we would use the background space-time on which the scalar moves, obtaining quantum field theory on a given, possibly curved background. But what do we do if the metric itself is one of the fields to be quantized? If we use it to smear itself, the resulting object becomes ugly, non-linear and too complicated as one of the basic quantities of a quantum field theory. If we introduce a separate metric just for the purpose of smearing the dynamical fields of gravity, this auxiliary input is likely to remain in the results derived from the theory. We would be quantizing gravitational excitations on a given space-time background, not the full gravitational field or space-time itself. We would be violating the great insight of general relativity by formulating physics on an auxiliary space-time, rather than realizing gravity as the manifestation of space-time geometry.

Fortunately, it is possible to smear fields and yet avoid the introduction of auxiliary metrics. Back to the scalar, we can choose to smear p_ϕ instead of ϕ. The momentum of a scalar field is a scalar density; it transforms under coordinate transformations with an extra factor of the Jacobian for the coordinate change. This is a direct consequence of the definition $p_\phi = \partial\mathcal{L}/\partial\dot{\phi}$ as a derivative of the Lagrangian density. Explicitly, the canonical variable $p_\phi = \sqrt{\det q}\,\dot{\phi}$ already carries the correct measure factor which need not be introduced by an auxiliary metric.[6] The smeared version $p_\phi[\lambda] := \int d^3y \lambda(y) p_\phi(y)$ is well-defined for any function λ, and it suffices to remove delta functions from the Poisson algebra: $\{\phi(x), p_\phi[\lambda]\} = \lambda(x)$.

[6] Although $\sqrt{\det q}$ appears in the relationship between p_ϕ and $\dot{\phi}$, from the viewpoint of Poisson geometry p_ϕ (but not $\dot{\phi}$) is independent of the metric: $\{p_\phi, p^{ab}\} = 0$ for the momenta p^{ab} of q_{ab}.

Holonomies and fluxes

For tensorial fields as we have them in gravity, background-independent smearings are often more difficult to find. Loop quantum gravity has provided suitable procedures for general relativity, but for this it must first transform from metric variables to connections with their conjugates, densitized triads. Connections and densitized vector fields turn out to have just the right transformation properties under coordinate changes that they can, with one loopy trick, be smeared background-independently. A well-defined quantization results, with several immediate implications for the basic operators encoding spatial geometry, as well as far-reaching and sometimes surprising consequences in the resulting dynamics.

Instead of using the spatial metric q_{ab}, spatial geometry is expressed by a densitized triad $E^a_i = \sqrt{\det q}\, e^a_i$ such that $E^a_i E^b_i = \det q\, q^{ab}$. The densitized triad is canonically conjugate to $K^i_a := K_{ab} e^{bi}$ in terms of extrinsic curvature K_{ab}. To obtain a connection with its useful transformation properties, we finally follow [2] and [13] to introduce the Ashtekar–Barbero connection $A^i_a = \Gamma^i_a + \gamma K^i_a$ with the spin connection Γ^i_a compatible with the densitized triad and a positive real number γ, the Barbero–Immirzi parameter (whose role for quantum geometry was realized in [74]).

The elementary objects loop quantum gravity takes for its representation are holonomies and fluxes,

$$h_e(A) = \mathcal{P}\exp\left(\int_e \mathrm{d}t\, \dot{e}^a A^i_a \tau_i\right) \quad \text{and} \quad F_S(E) = \int_S \mathrm{d}^2 y\, n_a E^a_i \tau_i\,,$$

first used in this context in [93]. A key advantage is that their algebra under Poisson brackets is well-defined, free of delta functions (unlike the algebra of fields), and yet independent of any background metric. Only the dynamical fields A^i_a and E^a_i are used, together with kinematical objects such as curves $e \subset \Sigma$ and surfaces $S \subset \Sigma$ as well as the tangent vectors $\dot{e}^a$ and co-normals n_a to them, but no independent metric structure.[7] All spatial geometrical properties are reconstructed from E^a_i via fluxes, and space-time geometry follows with A^i_a via holonomies once equations of motion (or rather the constraints of relativity) are imposed. In this way, loop quantum gravity provides a framework for background-independent quantum theories of gravity.

Once a well-defined algebra of basic objects has been chosen, one can determine its representations to arrive at possible quantum theories. In the connection representation, a complete set of states $\psi(A^i_a)$ is generated by holonomies as multiplication operators acting on $\psi(A^i_a) = 1$. In the case of loop quantum gravity, this has an immediate and general consequence. Holonomies take values in SU(2), and

[7] The co-normal, unlike the normal n^a, is metric-independent: for a surface $S\colon f = \text{const}$, $n_a = (\mathrm{d}f)_a$.

fluxes, depending on the momenta E^a_i, become derivative operators on SU(2) – just like angular momentum in quantum mechanics. For a dense set of states, only a finite number of holonomies (along curves intersecting the flux surface S) contribute to a given flux. Finite sums of angular momentum operators with discrete spectra have a discrete spectrum, too: spatial geometry is discrete; flux operators quantizing the densitized triad and thus encoding spatial geometry acquire discrete spectra. So do spatial geometrical quantities such as areas and volumes, as constructed in [3, 4, 95]. No extra assumptions are required; one merely has to fix the basic algebra and follow mathematical procedures to analyze its representations. A different algebra might lead to other properties, possibly not with discrete spatial geometry. But no alternative procedure providing a well-defined and smeared, yet background-independent quantization has been found. The holonomy–flux algebra suggested by its natural smearing, on the other hand, has a unique irreducible, cyclic, diffeomorphism-covariant representation, as proven by [67, 80]. Most of these properties are described by [98] in this volume.

Dynamics

Kinematical properties are elegant, simple, and largely unique. The theory starts to get considerably more messy when its dynamics is considered. Here, two major tasks must be performed. Dynamical operators, mainly the Hamiltonian constraint, must be defined from the basic ones, holonomies and fluxes. This is the constructive part of the task. For suitable constructions of the Hamiltonian constraint, the dynamics of the theory must then be evaluated, a process still full of many open issues.

At the constructive stage, many choices can be made, and strong consistency conditions must be respected. We are witnessing an epic battle between the liberating anarchy of choice and the uniformizing tyranny of constraints. Just writing a Hamiltonian constraint operator is tedious, but possible in many ways. There are ubiquitous factor-ordering ambiguities, as well as other choices specific to loop quantum gravity. On the other hand, the Hamiltonian constraint provides not only an equation of motion, it also generates a crucial part of the gauge transformations responsible for general covariance. A consistent quantization must keep gauge degrees of freedom as gauge, and not overly restrict the number of physical, non-gauge degrees of freedom. All this can usefully be formalized, as we will see later, providing strong algebraic conditions. They are so strong that to date, despite the abundant freedom of choices in constructing Hamiltonian constraint operators, no full consistent version has been found.

Once a consistent version of the dynamics exists, it must be evaluated. We must find solutions, and determine the physical observables they provide. Their values, finally, can be used for predictions such as small deviations from the expected classical behavior. Since no full consistent version has been found yet, and since all potential condidates are highly complicated, construction issues of the dynamics have so far dominated strongly over evaluation issues. In several model systems, on the other hand, the dynamics can often be simplified so much that it can be analyzed rather explicitly. Many useful techniques are now available, mainly in the context of effective equations. We will come back to these applications after discussing more details of the construction side.

Hamiltonian constraint

Specifically, the Hamiltonian constraint of general relativity in Ashtekar variables is

$$C[N] = \int_\Sigma \mathrm{d}^3x N \left(\varepsilon_{ijk} F^i_{ab} \frac{E^a_j E^b_k}{\sqrt{|\det E|}} - 2(1+\gamma^{-2}) K^i_a K^j_b \frac{E^{[a}_i E^{b]}_j}{\sqrt{|\det E|}} \right),$$

with the curvature F^i_{ab} of the Ashtekar connection and extrinsic curvature K^i_a. It has to vanish for all lapse functions $N(x)$, thus providing infinitely many constraints. If this is to be turned into an operator, using the basic expressions for holonomies and fluxes, several obstacles must be overcome.

First, there is the potentially singular inverse determinant of the densitized triad. No direct quantization exists since the densitized triad has been quantized to flux operators with discrete spectra, containing zero. Such operators lack densely defined inverses. Nevertheless, quantizations with the appropriate inverse as the semiclassical limit can be obtained making use of the classical identity

$$\left\{ A^i_a, \int \sqrt{|\det E|} \mathrm{d}^3x \right\} \propto \varepsilon^{ijk} \varepsilon_{abc} \frac{E^b_j E^c_k}{\sqrt{|\det E|}} \tag{11.14}$$

(or variants) as introduced by [103]. There is no inverse on the left-hand side. Instead, the expression involving the densitized triad is the spatial volume which (with some regularization) can be quantized directly. The connection components can be expressed via holonomies, and the Poisson bracket will, at the quantum level, be quantized to a commutator divided by $i\hbar$. A well-defined operator results with the right-hand side of (11.14) as the desired semiclassical limit by construction.

The remaining factors in the Hamiltonian constraint involve the Ashtekar curvature as well as extrinsic curvature. For the curvature components F^i_{ab}, we can use

$$s^a_1 s^b_2 F^i_{ab}(x) \tau_i = \Delta^{-1}(h_\lambda - 1) + O(\Delta) \tag{11.15}$$

where λ is a small loop starting at a point x, spanning a coordinate area Δ, and with tangent vectors s_1^a and s_2^a at x. On the right-hand side, the holonomy h_λ can readily be quantized, and to leading order provides curvature components as required for the constraint.

Extrinsic curvature, finally, is a more complicated object in terms of the basic ones but can be obtained from what has been provided so far:

$$K_a^i \propto \left\{ A_a^i, \left\{ \int \mathrm{d}^3x F_{ab}^i \frac{\varepsilon^{ijk} E_j^a E_k^b}{\sqrt{|\det E|}}, \int \sqrt{|\det E|} \mathrm{d}^3x \right\} \right\}$$

expresses extrinsic curvature in terms of a nested Poisson bracket involving the spatial volume and the first term of the constraint, already provided by the preceding steps.

In this way, holonomy and flux operators make up the Hamiltonian constraint operator $\hat{C}$ as a densely defined operator (including, in the non-vacuum case, regular matter Hamiltonians again exploiting (11.14) following [104]). It determines the physical solution space by its kernel: physical states $\psi(A)$, assumed again in the connection representation, must satisfy $\hat{C}\psi(A) = 0$, or $(\hat{C} + 8\pi G \hat{H}_{\text{matter}})\psi(A, \phi, \ldots) = 0$ if matter is present. The action of the constraint is rather complicated, as visualized schematically in Figure 11.1.

Physical states of interest normally belong to zero in the continuous part of the spectrum of $\hat{C}$, which requires the introduction of a new physical Hilbert space spanned by the solutions to the quantum constraint equation and equipped with a suitable physical inner product. Constructing the physical Hilbert space leads over

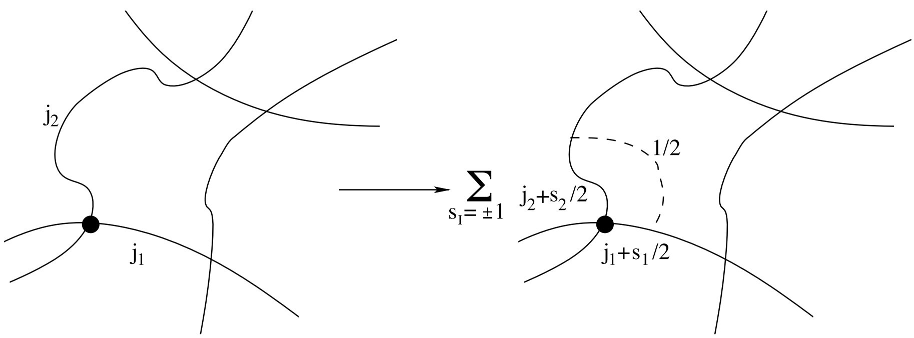


Figure 11.1 Schematically, the local action of the Hamiltonian constraint operator on a state. States are generated by holonomies as multiplication operators, visualized by the graph formed by all curves e used. Moreover, labels j_e pick matrix elements of SU(2)-valued $h_e(A)$. Due to (11.15), new curves and vertices are typically created when the Hamiltonian constraint acts.

to evaluating the dynamics, for physical states would provide predictions from expectation values of observables. But this stage has been brought to completion only in a few simple (and very special) models; in general, the outlook toward a full implementation is rather pessimistic.

Here, effective constraint techniques become useful because they allow one to address physical properties, corresponding to observables in the physical Hilbert space, without having to deal explicitly with states. Conditions from the physical inner product are rather implemented by reality conditions for expectation values and quantum variables such as fluctuations, which can be done more simply and more generally than for entire states. Effective constraints thus provide a good handle on generic properties of physical observables, at least in semiclassical regimes. Especially in cosmological situations they provide an ideal framework. In inhomogeneous contexts, they allow a detailed discussion of the anomaly issue, and show whether the different effects expected from the basic operators of quantum gravity can lead to a consistent form of the dynamics.

Consistency conditions in the presence of ambiguities are useful, for they constrain the choices. But the question remains whether loop quantum gravity can be fully consistent at all – implementing covariance at the quantum level might, after all, not leave any consistent physical states. If the consistency conditions are weak, on the other hand, ambiguities would remain even at the physical level. In between, at a razor-thin balance between ambiguity and constraints, lies the case of a unique consistent theory, a possibility which exists in loop quantum gravity but for which at present no evidence has been found.

In model systems it has at least been shown that consistency in the form of covariance can be achieved even in the presence of quantum corrections resulting from the discreteness. This statement may come as a surprise, for discrete spatial or space-time structures are naively expected to break local Lorentz or even rotational symmetries. Nevertheless, at the level of effective equations one can see that covariance can be respected – but it cannot leave the classical algebra of local symmetries invariant. While quantum gravity corrections provide a consistent deformation of general relativity – preserving the number of independent gauge generators – the algebra is truly deformed. Gauge transformations are no longer local Lorentz or coordinate transformations, but of a different type relevant for the quantum space-time structures derived from loop quantum gravity.

Much of these conclusions make use of results obtained through several years for models of loop quantum cosmology. We will now start a general exposition beginning with the simplest isotropic models, introducing the key features (and sources of ambiguities), and then leading over to the discussion of covariance in the next section.

11.3.2 Isotropic loop quantum cosmology

Loop quantum cosmology provides quantizations of symmetry-reduced models of general relativity, starting with isotropic ones in the simplest cases. It is thus a mini-superspace quantization, but not only that. There are explicit relationsips between the states and the algebra of basic operators in a model of loop quantum cosmology, which shows how they descend from the analogous expressions in the full theory. For different kinds of reductions such as implementing isotropy, homogeneity or spherical symmetry, this has been constructed in [23, 24, 43]. In particular, properties of the basic operators, for instance the discreteness of the flux spectra, are preserved and realized also in cosmological models. Crucial implications for the discreteness of spatial geometry can then be analyzed in their dynamical context. Qualitatively, all applications of loop quantum cosmology rely on this preservation of discreteness properties by the symmetry reduction introduced in [36].

While existing methods do not allow one to derive the dynamics of models directly from full constraint operators, it can be constructed in an analogous way using all the construction steps sketched for the full constraint. Also here, crucial properties are preserved. And even though the incompletely known relationship to the full theory introduces additional ambiguities in the dynamics of models, reliable conclusions can still be drawn provided the considerations are generic enough. Here, the need for sufficiently general parameterizations of ambiguities arises.

Basic variables

In spatially flat isotropic models, the basic canonical fields reduce to $A_a^i = \widetilde{c}\delta_a^i$ and $E_i^a = \widetilde{p}\delta_i^a$ with two dynamical variables $\widetilde{c}$ and $\widetilde{p}$. Relating the densitized triad to the spatial metric shows that $|\widetilde{p}| = a^2$ is given by the scale factor. (Due to the freedom of choosing an orientation of the triad, $\widetilde{p}$ can take both signs.) Classically, $\widetilde{c} = \gamma\dot{a}$ for spatially flat models.

To define the Poisson structure, we pull back the full symplectic form $(8\pi\gamma G)^{-1}\int \mathrm{d}^3x\delta A_a^i \wedge \delta E_i^a$ by the embedding $(\widetilde{c},\widetilde{p}) \mapsto (A_a^i, E_i^a) = (\widetilde{c}\delta_a^i, \widetilde{p}\delta_i^a)$. Due to homogeneity, the resulting integral diverges if we integrate over all of space and space is infinite, but homogeneity also implies that we may just integrate over any finite chunk of coordinate volume V_0, say, and still get the complete reduced symplectic form, $3V_0(8\pi\gamma G)^{-1}\mathrm{d}\widetilde{c}\wedge\mathrm{d}\widetilde{p}$. It depends on the arbitrary V_0, which can be hidden in the canonical variables by redefining them as

$$c = V_0^{1/3}\widetilde{c} \quad \text{and} \quad p = V_0^{2/3}\widetilde{p}.$$

In the end, physical quantities must be ensured to depend only on combinations of c, p and possibly other ingredients such that they are insensitive to changes of V_0. (Changing V_0 is independent of changing coordinates, thereby rescaling the scale factor. While $\tilde{c}$ and $\tilde{p}$ depend on the scaling but not on V_0, c and p depend on V_0 but not on the scaling.)

Isotropic connections and densitized triads result in specific versions of holonomies and fluxes. They take especially simple forms when evaluated along curves or surfaces making use of structures provided by the homogeneity group, i.e. curves along generators of translations and surfaces transversal to them. In a homogeneous context, fluxes are simply the triad components multiplied by the coordinate area of the surface, and holonomies can be computed explicitly: $h_{e_j}(A_a^i) = \exp(\mu c \tau_j)$ with $\mu = \ell_0 / V_0^{1/3}$ for a curve e_j of coordinate length ℓ_0 along the direction $\dot{e}_j^a = (\partial/\partial x^j)^a$ in Cartesian coordinates. If we take surfaces for fluxes of edge length the same size as e_j, fluxes are $F_{S_j}(E_i^a) = \mu^2 p \tau_j$ for a surface S_j transversal to $(\partial/\partial x^j)^a$.

These curves and surfaces are particularly useful because they can be arranged in a regular lattice of spacing ℓ_0. While such a setting is not the most general one (and other curves would lead to different expressions for holonomies as computed for instance in [53]), it allows one to capture all the degrees of freedom in a homogeneous model. More importantly, it allows for rough contact with the full theory, for the regular lattices envisioned here can be thought of as lattices for states in the full theory. In particular, for comparisons of actions of the dynamical operators, such a relationship is convenient. It also allows us to find a useful interpretation of the parameter μ characterizing isotropic holonomies and fluxes: for a regular lattice of spacing ℓ_0, contained in a region of size V_0, $\mu^3 = \ell_0^3 / V_0 =: \mathcal{N}^{-1}$ is the inverse number of lattice sites. This parameter can be seen as providing information about an underlying discrete inhomogeneous state which appears even in a homogeneous reduction. In fact, properties of $\mathcal{N}$ such as its size or possible dynamical features play a crucial role in Hamiltonian quantum evolution.

Representation

From isotropic holonomies and fluxes, we can construct their loop representation by analogy with the full theory. States in the connection representation are functionals of holonomies and are thus superpositions of the basic states $|\mu\rangle$, or $\exp(i\mu c)$ as functions of the isotropic connection component. Here, μ can be an arbitary real parameter. An inner product of states can be derived from integration theory on spaces of connections, as developed for instance in [6]. For isotropic models, this makes all states $|\mu\rangle$ orthonormal.

Upon completion to a Hilbert space, a general state takes the form of a countable superposition of $\exp(i\mu c)$ with $\mu \in \mathbb{R}$:

$$\psi(c) = \sum_{I \in \mathcal{I} \subset \mathbb{R}, \text{countable}} f_I \exp(i\mu_I c) \tag{11.16}$$

such that $\sum_{I\in\mathcal{I}} |f_I|^2$ exists. These states form a Hilbert space equivalent to the ℓ^2-space formed by the normalizable sequences $(f_I)_{I\in\mathcal{I}}$ for all countable $\mathcal{I} \subset \mathbb{R}$. In this form, isotropic states result directly from the reduction of a full state, where all the holonomies reduce to exponentials of c with different exponents. Alternatively, the Hilbert space may be characterized as the space of square-integrable functions $\psi(c)$ on the Bohr compactification of the real line, a compact space containing the real line densely and equipped with the measure

$$\int \mathrm{d}\nu(c)\psi(c) = \lim_{C\to\infty} \frac{1}{2C} \int_{-C}^{C} \psi(c)\mathrm{d}c$$

using the normal Lebesgue measure on the right-hand side. This Hilbert space is non-separable.

Basic operators act by multiplication or differentiation:

$$\widehat{\exp(i\delta c)}|\mu\rangle = |\mu + \delta\rangle \tag{11.17}$$

$$\hat{p}|\mu\rangle = \frac{8\pi}{3}\gamma \ell_{\mathrm{P}}^2 \mu |\mu\rangle, \tag{11.18}$$

as indeed follows from the unique holonomy–flux algebra in the full theory as an induced representation. A Wheeler–DeWitt representation, by contrast, would not be related to such a representation of the full theory.

Properties of the loop representation are markedly different from the Wheeler–DeWitt one, but they closely mimick properties of the full holonomy–flux representation:

- There is a discrete spectrum of $\hat{p}$. While there is a continuous range for μ, all eigenstates $|\mu\rangle$ are normalizable. For a non-separable Hilbert space such as we are dealing with here, normalizability of eigenstates does not imply that the eigenvalues form a countable subset of the real line. In such a situation, the normalizability condition is more general, for it is insensitive to what topology one uses on the set of eigenvalues. Even if any real number can appear as an eigenvalue, this would form a discrete set if one uses a discrete topology of the real line (for instance, one where every subset is an open neighborhood). Normalizability of eigenstates will also be one of the crucial properties for consequences of flux spectra in loop quantum cosmology.

- There is no operator for c, and only holonomies are represented. Trying to derive an operator for c from holonomies, for instance by taking a derivative of the action of $\widehat{\exp(i\delta c)}$ by δ at $\delta = 0$, fails because holonomies are not represented continuously in δ: $\langle\mu|\widehat{\exp(i\delta c)}|\mu\rangle = \langle\mu|\mu+\delta\rangle = \delta_{\delta,0}$ is not continuous.

These and other properties of the Bohr compactification as used in loop quantum cosmology are discussed in [8, 66]; for the case of a quantum cosmology based on ADM variables, see [73].

Since these are the same basic properties as realized for the full holonomy–flux algebra, they have the same qualitative implications for the dynamics. Once used for the construction of operators such as the Hamiltonian constraint, they lead to specific quantum geometry corrections in loop quantum cosmology as they do in the full theory. Specifically:

1. Only almost periodic functions of c are represented as operators on states (11.16), and must be expressible as $\psi(c) = \sum_{I\in\mathcal{I}\subset\mathbb{R},\text{countable}} f_I \exp(i\delta_I c)$. The Hilbert space does not allow an action of c on its dense subset of triad eigenstates; any appearance of c not of almost periodic form, such as the polynomial in the isotropic Hamiltonian constraint, must be expressed in terms of almost periodic functions by adding suitable higher-order corrections in c. They become significant for large values of the curvature, via $\ell_0\widetilde{c} = c/\mathcal{N}^{1/3}$ for a regular distribution of edges where $\delta_I \sim \mathcal{N}^{-1/3}$.
2. The isotropic flux operator $\hat{p}$ has a discrete spectrum containing zero and so lacks a direct, densely defined inverse. Well-defined versions can be obtained via identities such as

$$\frac{i}{\delta}e^{i\delta c}\{e^{-i\delta c}, |p|^{r/2}\} = \{c, |p|^{r/2}\} = \frac{4\pi\gamma G r}{3}|p|^{r/2-1}\,\mathrm{sgn}\,p \tag{11.19}$$

   which mimick the crucial one (11.14) used in the full theory. For $0 < r < 2$ we are expressing an inverse of p on the right-hand side, but do not need an inverse on the left-hand side; well-defined and even bounded operators for inverse powers of p result, as derived in [19]. Within this range, r is unrestricted by general considerations; it thus appears as an ambiguity parameter (see [22] for further discussions of ambiguities). Also here, corrections to classical expressions arise, in this case for small flux values $\ell_0^2\widetilde{p} = p/\mathcal{N}^{2/3}$ near the Planck scale. From a quantization of the left-hand side of (11.19), eigenvalues can readily be derived. They have the form $\frac{4}{3}\pi\gamma G r|p|^{r/2-1}\alpha_r(p)$, where $\alpha_r(p) \sim 1$ asymptotically for large $p/\mathcal{N}^{2/3} \gg \ell_\mathrm{P}^2$ while $\alpha_r(p) \to 0$ rapidly for $p \to 0$, cutting off the divergence of $|p|^{r/2-1}$. The correct semiclassical limit is guaranteed by α_r approaching one, which it does from above irrespective of quantization ambiguities.

Difference equation

The main quantity for which these considerations play a role is the Hamiltonian constraint, in isotropic variables providing the Friedmann equation

$$\frac{c^2\sqrt{|p|}}{\gamma^2} = \frac{8\pi G}{3} H_{\text{matter}}$$

with the matter Hamiltonian $H_{\text{matter}} = V_0 a^3 \rho$. Classically, its dependence on c is via c^2, which is not almost periodic. There are many ways to express c^2 in terms of almost periodic functions such that they approximate c^2 for small curvature, $c \ll 1$. In general terms, properties of the quantum representation space require the replacement of c^2 to be a bounded function of c, given by a normalizable superposition $\exp(\delta_I c)$ with δ_I in a countable subset of the real line. Each exponential acts by a shift of triad eigenvalues, rather than a derivative operator, and so, following [20, 21], upon loop quantization the Hamiltonian constraint equation becomes a difference equation for components of a wave function in the triad representation.

A choice usually made is $c^2 \sim \delta^{-2} \sin^2(\delta c)$, but others are possible, constituting further ambiguities. A Hamiltonian constraint operator

$$\hat{C}_{\text{iso}} = \frac{3}{8\pi G\gamma^2\delta^2} \widehat{\sin\delta c}^{\,2} \widehat{\sqrt{|p|}} - \hat{H}_{\text{matter}}$$

results. Expanding a state $|\psi\rangle = \sum_\mu \psi_\mu |\mu\rangle$ in the triad eigenbasis $\{|\mu\rangle\}_{\mu\in\mathbb{R}}$, $\hat{C}_{\text{iso}}|\psi\rangle = 0$ is equivalent to the difference equation

$$C(\mu+2\delta)\psi_{\mu+2\delta} - 2C(\mu)\psi_\mu + C(\mu-2\delta)\psi_{\mu-2\delta} = \frac{8\pi G}{3}\gamma^2\delta^2 \hat{H}_{\text{matter}}(\mu)\psi_\mu \tag{11.20}$$

as derived in [21], where $C(\mu)$ are eigenvalues of $\widehat{\sqrt{|p|}}$, e.g. proportional to $\sqrt{|\mu|}$. Additional matter fields have been suppressed in the notation, which would provide further independent variables in ψ_μ acted on by the matter Hamiltonian $\hat{H}_{\text{matter}}(\mu)$. (Unless there are curvature couplings, matter Hamiltonians depend on the densitized triad but not on the connection; the right side of (11.20) is then not a difference expression. Fermions, which require a coupling to the gravitational connection, have been discussed in this context in [32], and non-minimally coupled scalars in [35].)

Due to the inequivalence of representations, this difference equation replaces the differential Wheeler–DeWitt equation. At this dynamical level, after choosing the representation of c^2 by almost periodic functions, the dynamics can be restricted to a separable subsector of the initial non-separable Hilbert space. From this perspective, one could have started directly from a separable Hilbert space by choosing a specific countable set $\{\mu_I\} \subset \mathbb{R}$, such as $\mu_I = I\mu_0$ with integer I as originally defined in [21].

Many dynamical properties would follow in the same way since non-separable features appear only at the kinematical level.

However, there is a good reason for dealing with a non-separable kinematical Hilbert space, allowing for all real values of μ. To capture the most general and viable dynamics, one expects lattice refinement to happen, with a discreteness scale changing dynamically. Indeed, full Hamiltonian constraint operators create new edges and vertices, changing the graph on which a state is defined, as illustrated in Figure 11.1. In an isotropic context, a trace of this feature must still be left since the dynamical parameter δ came from the edge length ℓ_0 used in elementary holonomies. While V_0 in $\delta = \ell_0/V_0^{1/3}$ must remain constant as a mere auxiliary parameter introduced in the setup, ℓ_0 must be adapted to the geometry. Geometrical distances, after all, increase as the universe expands and the region of coordinate size V_0 has a growing geometrical size $V_0 a^3$. If a lattice state for a larger region has more sites, as suggested by the creation of vertices, ℓ_0 must shrink if a grows; otherwise the discrete size $a\ell_0$ would be blown up macroscopic. Here, the evolutionary aspects capture expected properties of full physical states, solving a vertex-creating Hamiltonian constraint, arranged by volume eigenvalues as internal time. (From a more general perspective, dynamical lattices and their role in quantum gravity have been discussed in [75, 106, 107].) Then, also $\delta(\mu)$ must not be constant but rather a decreasing function. (Just as in (11.15), for the regularization of the isotropic F^i_{ab}, or the c^2 in the Friedmann equation, by $\delta^{-2}\sin^2(\delta c)$, we use the coordinate area ℓ_0^2 of loops, not the geometrical area $\ell_0^2 a^2$ which might well be constant as a changes.)

In summary, lattice refinement means that δ depends on the size of the universe, and $\delta(\mu)$ is not a constant. The resulting difference equations, as most generally formulated for anisotropic models in [46], are not equidistant and do not allow simple restrictions to separable sectors. A large class of different dynamics can thus be formulated in each single model.[8] Consistency conditions do exist restricting the freedom even in homogeneous models, but systematic investigations have only just begun, as e.g. in [51]. While specific details of the difference equations of loop quantum cosmology are not yet fully determined, owing to quantization ambiguities, there are several key generic features implying further properties. They are brought out most clearly in solvable models.

Harmonic cosmology

Sometimes, solutions to the difference equations of loop quantum cosmology, which are linear but with non-constant coefficients, can be found. They may be studied

[8] Several improvised models have been formulated and sometimes analyzed in detail, based on further ad-hoc assumptions to constrain ambiguities. Since this cannot capture the generic behavior, specific details of the results are unlikely to be reliable.

numerically, and lead to interesting issues – especially in cases where they are not equidistant, as investigated so far in [87,97]. But in all cases, computing observables from the resulting wave functions and arriving at sufficiently generic conclusions is challenging.

An effective treatment at this stage becomes much more powerful. Initially, one may expect technical difficulties due to the non-linear and non-polynomial nature of the Hamiltonian constraint obtained after the loop replacement of c^2. Fortunately, isotropic loop quantum cosmology offers an exactly solvable system where, in a specific factor ordering of the constraint operator, the dynamics is essentially free. From the point of view of quantum dynamics, this model is closely related to the harmonic oscillator of quantum mechanics, which forms the basis of most of the perturbation theory framework for interacting quantum field theories. Similarly, the harmonic model of loop quantum cosmology plays a crucial role for perturbations in quantum gravity.

Specifically, according to [25, 26], we obtain the solvable model for spatially flat isotropic loop quantum cosmology with a free, massless scalar, whose loop-modified Hamiltonian constraint is

$$\frac{4\pi G}{3}\frac{p_\phi^2}{p^{3/2}} = \sqrt{p}\gamma^{-2}\delta(p)^{-2}\sin^2(\delta(p)c) \sim \sqrt{p}c^2 + O(c^4)\,.$$

Here, $\delta(p)$ (or $\mathcal{N}(p) = \delta(p)^{-3}$) is allowed to depend on the triad, incorporating different refinement schemes of the discrete structure. For power laws $\delta(p) \propto p^x$, we introduce $V = p^{1-x}/(1-x)$ and $J = p^{1-x}\exp(ip^x c)$ as non-canonical basic variables forming an $\mathrm{sl}(2,\mathbb{R})$ algebra

$$[\hat{V},\hat{J}] = \hbar\hat{J}, \quad [\hat{V},\hat{J}^\dagger] = -\hbar\hat{J}^\dagger, \quad [\hat{J},\hat{J}^\dagger] = -2\hbar\hat{V}$$

upon quantization. Thanks to the free scalar, the dynamics is controlled by the deparameterized Hamiltonian in ϕ, $\hat{p}_\phi \propto |\frac{1}{2i}(\hat{J} - \hat{J}^\dagger)| =: \hat{H}$.

This Hamiltonian is linear for all x, which for our linear algebra of non-canonical basic variables implies solvability. (Again, the absolute value is irrelevant for effective equations.) Following our general analysis of effective descriptions, we can jump directly to equations of motion for expectation values of the basic variables. They are coupled to each other, but not to quantum variables such as fluctuations, as normally occurs in interacting quantum systems. Instead of a non-linear set of infinitely many differential equations, we have a finitely coupled set of linear equations. For the basic variables, we have

$$\frac{\mathrm{d}\langle\hat{V}\rangle}{\mathrm{d}\phi} = \frac{\langle[\hat{V},\hat{H}]\rangle}{i\hbar} = -\tfrac{1}{2}(\langle\hat{J}\rangle + \langle\hat{J}^\dagger\rangle)$$
$$\frac{\mathrm{d}\langle\hat{J}\rangle}{\mathrm{d}\phi} = \frac{\langle[\hat{J},\hat{H}]\rangle}{i\hbar} = -\tfrac{1}{2}\langle\hat{V}\rangle = \frac{\mathrm{d}\langle\hat{J}^\dagger\rangle}{\mathrm{d}\phi}.$$

As a difference from the previous discussion of effective systems, we now have to take into account the complex nature of our variable J. Reality conditions are required for physical results, which would be implemented by the physical inner product in a Hilbert space representation. Here, we are not using states but work directly with expectation values and moments. The adjointness relation $\hat{J}\hat{J}^\dagger = \hat{V}^2$ for our basic operators implies, upon taking an expectation value, a reality condition relating $|\langle\hat{J}\rangle|^2 - \langle\hat{V}\rangle^2$ to moments of $\hat{V}$ and $\hat{J}$. It turns out that the specific combination of moments involved is constant in ϕ-evolution, and that it is of the order $\hbar$ for a semiclassical state. (For more details, see [25,29].) Just requiring that the state is semiclassical only once, for instance at large volume, ensures that the reality condition reads $|\langle\hat{J}\rangle|^2 - \langle\hat{V}\rangle^2 = O(\hbar)$ at all times. With this condition, the general solution is

$$\langle\hat{V}\rangle(\phi) = \langle\hat{H}\rangle\cosh(\phi-\delta), \quad \langle\hat{J}\rangle(\phi) = -\langle\hat{H}\rangle(\sinh(\phi-\delta) - i)$$

with the conserved $\langle\hat{H}\rangle$ and another integration constant δ. For a large class of states we thus have an exact realization of a bounce, with the volume bounded away from zero. (Effective equations thus easily confirm the numerical results of [9], at least as far as physical expectation values are concerned. For the dynamical behavior of moments, the generic treatment of [27, 29] based on effective equations shows crucial differences from the numerics done for only a specific class of states.)

However, the system is harmonic, and so its behavior is not easily generalizable to realistic models. While it is interesting that bounce models can be derived in this way, taking this at face value without realizing the solvable nature of the model would provide a view with severe limitations. After all, the harmonic oscillator does not provide general insights into quantum dynamics, and free quantum field theories allow no glimpse at the rich features of interacting ones. Only a systematic analysis of perturbations around the free model, starting with interaction terms and then introducing inhomogeneities, can provide a clear picture.

Quantum Friedmann equation

As already briefly discussed in the context of Wheeler–DeWitt quantum cosmology, effective equations of quantum cosmology are necessarily of higher dimension than the classical equations. Quantum degrees of freedom such as fluctuations couple to expectation values in non-solvable models, and no adiabatic or other regime has been determined where quantum variables could be expressed completely in terms of effective contributions depending only on the expectation values. Quantum variables are true degrees of freedom, of significance for the dynamics. Without a clear vacuum state to expand around, moreover, suitable states are less restricted than they are for low-energy effective actions. Via initial values for moment equations, this leaves the state dependence as a crucial contribution of equations, to be taken into account for sufficiently general conclusions. Since not much is known about details of the quantum state of the universe, results should be sufficiently insensitive to its properties.

Perturbation theory by the moments, including them order by order, is required for any model which is not harmonic. Examples are models with non-vanishing spatial curvature or a cosmological constant, as analyzed in [41] (or numerically in [16]). Another class of important models is that where the scalar is no longer free, or at least acquires a mass term. We are then dealing with a time-dependent Hamiltonian in the ϕ-evolution, and ϕ may not even provide a good internal time if the potential leads to turning points of $\phi(\tau)$. At this stage, a more general analysis is required which is now available in the form of effective constraints as per [42,48]: we do not have to deparameterize the system before quantizing it or computing effective equations. We can compute effective constraints directly and analyze them to find the physical quantum phase space. As one of the results, we can apply the deparameterized framework even if there is a potential, provided it changes sufficiently slowly. This is exactly the case of interest for inflationary or other early-universe cosmology, and so we can derive effective equations for these situations.

In the presence of a potential, the Friedmann equation receives the following quantum corrections:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\left(\rho\left(1-\frac{\rho_Q}{\rho_{\mathrm{crit}}}\right) \pm \frac{1}{2}\sqrt{1-\frac{\rho_Q}{\rho_{\mathrm{crit}}}}\,\eta(\rho-P) + \left(\frac{\rho-P}{\rho+P}\right)^2\eta^2\right),$$

as derived in [29,31]. (The top line, corresponding only to the solvable model with a free scalar, was obtained earlier in [100,101]. It sums up all higher-order corrections due to holonomies, as expanded explicitly in [12]. In the usual terminology of quantum field theory, it represents the tree-level approximation since its main form is independent of quantum back-reaction.) Here, P is pressure, η parameterizes quantum correlations, and

$$\rho_Q := \rho + \varepsilon_0\rho_{\mathrm{crit}} + (\rho - P)\sum_{k=0}^{\infty}\varepsilon_{k+1}\left(\frac{\rho-P}{\rho+P}\right)^k$$

is a quantum-corrected energy density with fluctuation parameters ε_k. The critical density $\rho_{\rm crit} = 3/8\pi G\gamma^2\delta^2 V_0^{2/3} a^2$ results from the loop quantization, bringing in the scale $\delta(a)$ whose form is determined by lattice refinement.

This equation contains all quantum variables in η and ρ_Q, subject to their own dynamics. Understanding the behavior of generic universe models in loop quantum cosmology requires a high-dimensional dynamical system to be analyzed. Only a few simple cases allow strict conclusions about the presence of a bounce. If $\rho = P$ (a free, massless scalar), we recover the solvable scenario. For $\rho + P$ large (large p_ϕ, i.e. kinetic domination), all corrections involving quantum variables are suppressed; at least for a certain amount of time, the behavior can be expected to be close to the solvable one.

In general, however, quantum variables may yank the universe away from the simple bouncing behavior of the solvable model. While the fundamental dynamics based on difference equations for states remains non-singular, following [18], the general effective one is still unclear. Volume expectation values could approach a non-zero constant asymptotically without bouncing back to large values, a picture which might resemble that of the emergent universe of [64,65]. Most likely, generic evolution would bring us to a highly quantum regime where effective equations or simple near-solvable models break down. Ultimately, the fate of a universe can be studied only based on the fundamental difference equations, but effective equations can show well how such severe states are approached.

The interplay of different quantum corrections

Loop quantum cosmology leads to different types of quantum corrections in its effective equations: holonomy corrections, inverse triad corrections, and quantum back-reaction. Depending on parameters, they may play different roles in any given regime, and some of them might dominate. What exactly happens can be determined only with a sufficiently general parameterization of correction terms. Here, all ambiguities and effects such as lattice refinement must be considered.

Especially holonomy and inverse triad corrections are often closely related to each other: they both result from the discrete geometry, although in different ways. Comparing them can thus provide more restrictions on free parameters of the full theory than any one of the corrections would allow individually. They are both related to the size of elementary building blocks of a discrete geometry. In a nearly isotropic distribution in a region of coordinate size V_0, the classical volume on the left-hand side of

$$V_0 a(\phi)^3 = \mathcal{N}(\phi) v(\phi)$$

is replaced by the right-hand side in the discrete picture. This elementary relationship first tells us that, dynamically, there are two free functions in the discrete picture for one free function $a(\phi)$ in the continuum picture: the size $v(\phi)$ and number $\mathcal{N}(\phi)$ of discrete sites. Both of them typically change in internal time: the lattice they form is being refined as time goes on.

As derived earlier, one of these parameters, v, enters the basic expressions for holonomies

$$\exp(i\ell_0 c/V_0^{1/3}) = \exp(ic/\mathcal{N}^{1/3}) = \exp(i\gamma v^{1/3}\dot{a}/a)$$

using $c = V_0^{1/3}\tilde{c} = V_0^{1/3}\gamma\dot{a}$. Inverse triad corrections in isotropic models provide a correction factor α depending on $p/\mathcal{N}^{2/3} = (V_0/\mathcal{N})^{2/3}a^2 = v^{2/3}$. For $v^{1/3} \sim \ell_{\mathrm{P}}$, inverse triad corrections differ strongly from the classically expected value $\alpha = 1$ (while holonomies remain nearly equal to the classically linear $\dot{a}/a$ for Hubble distances much larger than $v^{1/3}$). Here, quantum corrections can often be constrained. Since two different corrections depend on one parameter, their interplay can provide important synergistic effects in ruling out possible values, or even entire corrections.[9]

Phenomenologically, v shows up in the critical density of the effective Friedmann equation: $\rho_{\mathrm{crit}} = 3/8\pi G\gamma^2 v^{2/3}$. At this value, we would find the bounce of a universe sourced by a free, massless scalar, and even in other models its value affects details of the dynamics. This density is often assumed Planckian, which means that $v \sim \ell_{\mathrm{P}}^3$. (From black-hole entropy calculations, as developed in [7, 62, 76, 82], $\gamma \sim 0.2$.) But then, inverse triad corrections are large. If v is constant, which is also often assumed, inverse triad corrections would be large at all times, even at large total volume. A Planckian constant value of v is thus clearly ruled out, not by holonomy corrections alone but by a consistent combination with inverse triad effects. Further phenomenological analysis is in progress, by [15, 57, 71, 72, 83, 85, 86, 99]. Much stronger consistency conditions can be expected when inhomogeneities are included, to which we will turn next.

11.4 Consistent dynamics

A space-time covariant theory is a gauge theory with the gauge group given by space-time diffeomorphisms. As a consequence, the local conservation law $\nabla^\mu(G_{\mu\nu} - 8\pi G T_{\mu\nu})$ holds for the Einstein tensor $G_{\mu\nu}$ and the stress–energy tensor $T_{\mu\nu}$ of matter. Not all components of Einstein's equation are thus independent,

[9] We can also note that all quantum geometry corrections depend only on $v = V_0 a^3/\mathcal{N}$. As the size of discrete building blocks, it is insensitive to changing V_0 (both V_0 and $\mathcal{N}$ change proportionally to V_0) or coordinates ($V_0 a^3$ and $\mathcal{N}$ are scaling-invariant). Just like the classical equations, quantum corrections are invariant under these rescalings. For the scale factor, inverse triad corrections become significant at a characteristic scale $a_* = (\mathcal{N}/V_0)^{1/3}\ell_{\mathrm{P}}$ related to the Planck length. While a_* is not scaling-independent, it scales in the same way as a. Comparisons such as $a > a_*$ or $a < a_*$ to demarcate the classical and the strongly quantum regime are thus meaningful.

and some can be derived from the others. For consistency, effective equations of quantum gravity must show the same form of dependencies, or allow the same number of local conservation laws, or else degrees of freedom would be too much constrained to have a chance of providing the correct classical limit.

Consistency is automatically satisfied if one has space-time covariant higher-order corrections in Lagrangian form. At the Hamiltonian level, however, covariance is more difficult to check due to the lack of a direct action of space-time diffeomorphisms. There is rather a splitting into spatial diffeomorphisms acting as usual, and an independent generator deforming spatial constant-time slices within space-time. As in relativistic systems, all generators are constraints, taking center stage in a canonical analysis. The conservation law implies that the components $G_{0\nu} - 8\pi G T_{0\nu}$ of Einstein's equation are only of first order in time derivatives, since $\partial_0(G^0_\nu - 8\pi G T^0_\nu)$ must equal terms of at most second order in time. The same equation shows that these constraints are automatically preserved in time, given that $\partial_0(G^0_\nu - 8\pi G T^0_\nu)$ depends only on the Einstein equations at fixed time and their spatial derivatives. If Einstein's equation holds at one time, time derivatives of the components $G_{0\nu} - 8\pi G T_{0\nu}$ vanish; the constraints can consistently be imposed at all times. Also this condition must be preserved for effective equations. Otherwise, the system is overdetermined with more equations than unknowns and not enough consistent solutions would result.

From this perspective, the covariance condition takes the following form at the Hamiltonian level. For a generally covariant system, evolution in coordinate time is generated by the constraints, and the constraints must themselves be preserved in time. Thus, the gauge transformation $\delta_{\varepsilon^\mu} H[N^\nu] = \{H[N^\nu], H[\varepsilon^\mu]\}$ of a combination $H[N^\nu] := \int \mathrm{d}^3x N^\nu (G_{0\nu} - 8\pi G T_{0\nu})$ must vanish for all space-time vector fields N^ν, ε^μ when the constraints are satisfied. These Poisson bracket relations provide algebraic conditions for the constraints to be consistent, forming a so-called first-class system. Specifically for general relativity, we have

$$\{H[N^\mu], H[M^\nu]\} = H[K^\mu] \tag{11.21}$$

with $K^0 = \mathcal{L}_{M^a} N^0 - \mathcal{L}_{N^a} M^0$ and $K^a = \mathcal{L}_{M^a} N^a - \mathcal{L}_{N^a} M^a + q^{ab}(N^0 \partial_b M^0 - M^0 \partial_b N^0)$. It is a concise expression for space-time covariance, and is insensitive to the specific dynamics: the algebra is the same for all higher curvature actions (or with all kinds of matter), even though the specific constraint functionals do change. (A recent Hamiltonian analysis of higher curvature actions has been performed in [58], showing the constraint algebra explicitly.)

The canonical consistency requirement of an anomaly-free, covariant theory then states that the algebra of effective constraints must remain first class. As long as this algebraic condition is satisfied, the theory is completely consistent. The system of constraints must remain first class after including quantum corrections, but the specific algebra may change. Canonically, the realm of consistent theories is larger than at the Lagrangian level: while corrections to the action can be of higher curvature form, they all give rise to exactly the same constraint algebra. The Hamiltonian level, on the other hand, allows changes in the algebra as long as it remains first class; the theory may be consistently deformed. Thereby, changes to the quantum space-time structure can be captured. Whether or not non-trivial consistent deformations exist, and how their covariance can be interpreted, is a matter of analysis in specific quantum gravity models. We will illustrate the results of several models in loop quantum cosmology in the following section.

11.4.1 Cosmological perturbations

The general issues of constrained systems in gravity can easily be seen already for linear cosmological perturbations. For scalar modes, perturbing the lapse function as $N(t)(1+\phi(x,t))$ and the scale factor as $a(t)(1-\psi(x,t))$, and with a scalar matter source, the linearized components of Einstein's equation, in conformal time and longitudinal gauge, read

$$\partial_c\left(\dot{\psi}+\mathcal{H}\phi\right)=4\pi G\dot{\bar{\varphi}}\partial_c\delta\varphi \tag{11.22}$$

$$\nabla^2\psi-3\mathcal{H}\left(\dot{\psi}+\mathcal{H}\phi\right)=4\pi G\left(\dot{\bar{\varphi}}\delta\dot{\varphi}-\dot{\bar{\varphi}}^2\phi+a^2V_{,\varphi}(\bar{\varphi})\delta\varphi\right) \tag{11.23}$$

$$\ddot{\psi}+\mathcal{H}\left(2\dot{\psi}+\dot{\phi}\right)+\left(2\dot{\mathcal{H}}+\mathcal{H}^2\right)\phi=4\pi G\left(\dot{\bar{\varphi}}\delta\dot{\varphi}-a^2V_{,\varphi}(\bar{\varphi})\delta\varphi\right) \tag{11.24}$$

$$\partial_c\partial^c(\phi-\psi)=0 \tag{11.25}$$

with the conformal Hubble parameter $\mathcal{H}$. This set of equations is accompanied by the linearized Klein–Gordon equation for $\varphi=\bar{\varphi}+\delta\varphi$. The first two lines are of first order in time, and thus pose constraints for initial values. They are the linearized diffeomorphism constraint (11.22) and the Hamiltonian constraint (11.23). The next two lines are the diagonal and off-diagonal components of the spatial part of Einstein's equation, also linearized around Friedmann–Robertson–Walker. Using the background equations for $\mathcal{H}$ and the background scalar $\bar{\varphi}$, one can explicitly derive the Klein–Gordon equation from the rest. The system is thus overdetermined, but in a consistent way. Canonically, this follows from the fact that all equations result from a set of first-class constraints.

In these equations, perturbations ϕ, ψ and $\delta\varphi$ are used directly for the metric and the scalar. These quantities are subject to gauge transformations under changes of space-time coordinates, which also change the gauge. So far, the equations as written are in longitudinal gauge where the space-time metric remains diagonal under perturbations. More generally, scalar modes can be perturbed also in the off-diagonal part of the spatial metric by $\delta q_{ab}=\partial_a\partial_bE$, or in the space-time part by a perturbed shift vector $\delta N^a = \partial^a B$. Two new scalar perturbations E and B are introduced in addition to ϕ and ψ, and all four transform into each other under coordinate changes. For linear coordinate changes, however, the combinations

$$\Psi = \psi - \mathcal{H}(B - \dot{E}) \tag{11.26}$$

and

$$\Phi = \phi + \left(B - \dot{E}\right)^{\bullet} + \mathcal{H}(B - \dot{E}) \tag{11.27}$$

and a similar one for the matter perturbation remain unchanged, as specified by [14]; they form gauge-invariant observables of the linear theory. From a canonical perspective, these combinations are invariant under the flow generated by the constraints.

The ungauged evolution equations, containing E and B in addition to ϕ and ψ, can be expressed completely in terms of the gauge-invariant variables, as required on physical grounds. Here, the first-class nature of constraints is important, which ensures that the constraint equations, and thus (11.22) and (11.23), are gauge-invariant. Consistency as well as gauge invariance of the equations of motion is thus guaranteed by having a first-class algebra of constraints. Also here, the first-class nature serves as a complete requirement for covariance.

11.4.2 Gauges and frames

A first-class algebra of constraints ensures that physical evolution can be formulated fully in gauge-invariant terms. By the same property of the constrained system, frame-independence is also guaranteed: the gauge algebra corresponds to deformations of the spatial hypersurfaces given by a time function $t = \text{const}$. Gauge-invariant variables are insensitive to these deformations, and thus to the choice of time. But a time function (without specifying which spatial coordinates are to be held fixed) does not uniquely tell us how to take time derivatives for the equations of motion; this is only accomplished when we also specify a time evolution vector field t^a (such that $t^a \partial_a t = 1$). For a given time function, there are many choices for t^a, and the time evolution vector field may be changed independently of the foliation.

A fixed foliation of space-time into spatial slices with unit normals n^a allows us to express the freedom in choosing time evolution vector fields by the lapse function N and the shift vector N^a (with $N^a n_a = 0$) such that $t^a = N n^a + N^a$. These are exactly the functions which appear in the constraints generating evolution: for a fixed choice of t^a, or N and N^a, Hamiltonian equations of motion are $\dot{f} = \mathcal{L}_{t^a} f = \{f, H[N, N^a]\}$ for any phase-space function f. Since N and N^a appear as multipliers of first-class constraints, they can be chosen arbitrarily (except that we would like $N > 0$ for evolution toward the future). For a consistent set of first-class constraints, we can thus choose the frame freely, and different frame choices are guaranteed to produce consistent results. (In a space-time treatment, observables which are gauge as well as frame-independent can be derived following [52, 63]. In a reduced phase-space treatment, where one uses observables solving the classical constraints, one is working at a gauge-invariant, but not automatically frame-independent level.)

In classical relativity, cosmological perturbation equations can often be derived in much simper ways when a space-time gauge is chosen, such as the longitudinal gauge above or the uniform one where only matter fields are perturbed. Since gauge transformations are known to correspond to space-time coordinate transformations, one can verify directly that such gauges are possible. Moreover, choosing a gauge before deriving equations of motion from an action or Hamiltonian is equivalent to choosing a gauge in the general equations of motion. The situation is thus completely unambiguous.

When quantum effects are included, either in a full quantum theory or in an effective manner, the constraints change by quantum correction terms. Equations of motion change, as expected, and so do the form of gauge-invariant expressions, since it is the constraints which generate gauge transformations. In such a situation, gauge transformations and suitable gauge fixings can be analyzed only after the quantization has been performed and the corrected constraints are known. If gauge fixings are employed before quantization or before determining effective constraints, the choice of gauge fixing may not be compatible with the resulting corrected gauge transformations. Moreover, choosing different gauge fixings before doing the same kind of quantization would in general lead to different final results, making the procedure ambiguous even beyond unavoidable quantization ambiguities. Similarly, a reduced phase-space quantization is based on frame-fixing, although no gauge need be fixed.

For the different approaches used in relation to loop quantum gravity, several examples exist. Reference [55] develops methods to deal with a discretization possibly breaking gauge symmetries. Similar methods have then been used in [56] in spherically symmetric models with (partial) gauge fixing. Also the hybrid quantizations of Gowdy models in [81] rely on gauge fixing of the inhomogeneous generators. Reference [11] constructs discretized theories in three dimensions, respecting the space-time gauge, but [10] argues that this may not be possible in four space-time dimensions. References [78, 79] quantize two-dimensional parameterized models of field theories by the Dirac procedure and represent observables on the resulting physical Hilbert space. In this treatment, the discretization does not lead to inconsistencies but possibly to deformations of classical algebras of observables. Finally, reduced phase space methods fixing the frame by referring to an extra dust field are developed in [68, 69] for cosmological perturbations, and in [70] in spherical symmetry. Other treatments for cosmological perturbations are used, e.g., [1, 84] in gauge-fixed versions and [88] in a frame-fixed (reduced phase-space) way. Reference [96] provides a proposal by which cosmological perturbations, taking into account space-time discreteness, might be implementable by a consistent first-class algebra of constraints. So far, this has been realized for coupling two independent homogeneous patches.

The only valid treatment of a complicated gauge theory is by working without restrictions of the gauge throughout the quantization procedure, until the final gauge algebra resulting from the corrected constraints has been confirmed to be consistent. Here, the anomaly problem, confirming that a consistent deformation is realized, must be faced head-on and cannot be evaded. In the final equations one may choose one of the allowed gauges for further analysis, but gauges cannot be used to simplify the quantization. In the rest of this exposition, we follow these lines to illustrate the consistency of several effective sets of constraints incorporating some of the discreteness effects of loop quantum gravity.

11.5 Consistent effective discrete dynamics

If loop quantum gravity has a chance of being a viable quantum theory of gravity, the form of discrete quantum geometry it implies must give rise to effective dynamical equations satisfying the consistency conditions of a covariant theory. At the Hamiltonian level, this requires a first-class algebra of constraints. After the preparation in the preceding sections, we can now see what specific models indicate.

11.5.1 Constraint algebra

Consistent deformations implementing the effects of loop quantum gravity have been found in several different cases. Most of them use inverse triad corrections, which have been incorporated successfully in spherically symmetric models [37,50] as well as linear perturbations around spatially flat Friedmann–Robertson–Walker models [47, 49]. The situation for holonomy corrections is more restrictive; here, certain versions have been realized in spherically symmetric models [89] as well as for linear tensor and vector modes around Friedmann–Robertson–Walker models [33, 34]. However, so far no inhomogeneous model has been found where holonomy corrections in a complete form would be consistent. Here, the requirement of anomaly-freedom appears very restrictive.

Inverse triad corrections have been implemented consistently in several settings, and so are not ruled out by consistency considerations. In particular, corrections from the discreteness of quantum geometry are allowed and do not necessarily spoil covariance. The specific form of their implementation then tells us if and how space-time structures have to change due to quantum effects.

Based on formulas such as (11.14), inverse triad corrections arise for any term in the Hamiltonian constraint bearing components of the inverse densitized triad, such as $1/\sqrt{\det E^a_i}$. Reference [47] has consistently implemented these corrections for linear inhomogeneities, where the corrected constraint algebra is of the form (11.21) but with

$$K^a = \mathcal{L}_{M^a} N^a - \mathcal{L}_{N^a} M^a + \bar{\alpha}^2 \bar{N} a^{-1/2} \partial^a (\delta N_2 - \delta N_1)$$

while $K^0 = \mathcal{L}_{M^a} \delta N^0 - \mathcal{L}_{N^a} \delta M^0$ retains its classical form. In addition to the contribution $\bar{N} a^{-1/2} \partial^a (\delta N_2 - \delta N_1)$, which is expected classically for a linearization around Friedmann–Robertson–Walker models with $N_i = \bar{N} + \delta N_i$, there is the function $\bar{\alpha}$ (depending on the background scale factor a) arising from inverse triad corrections. An algebra of the same form arises for spherically symmetric models, with different matter couplings as discussed in [89]. The constraint algebra is anomaly-free: the system of constraints remains first class. But it is not exactly the classical algebra, and thus deformed. Inverse triad corrections from loop quantum gravity cannot amount to higher curvature corrections to the action since this would leave the constraint algebra unchanged. Rather, these corrections can only be understood as deforming local space-time symmetries while keeping covariance realized.

11.5.2 Cosmological perturbations

Consistent versions for quantum-corrected constraints allow one to analyze their implications for the dynamics. When constraints are corrected, not just evolution equations change but also the gauge transformations generated by the constraints. Thus, expressions for gauge-invariant observables depend differently on perturbations of the fields, which by itself may give rise to new effects. Other implications then follow from studying the dynamical evolution of gauge-invariant observables. (Some quantum gravity corrections have been implemented in gauge-fixed versions. They are formally consistent, but quite arbitrary in their implementation. For instance, it remains unclear how different gauge-fixings, all done before quantization, might affect the results. Moreover, some physical effects due to corrections to gauge-invariant observables can easily be overlooked.)

For linear perturbation equations around Friedmann–Robertson–Walker models, cosmological perturbation equations are the main application of consistent deformations. With a consistent deformation, perturbation equations form a closed set and can be written fully in terms of gauge-invariant variables. For inverse triad corrections, as derived in [49], they take the form

$$\partial_c \left(\dot{\Psi} + \mathcal{H}(1+f)\Phi\right) = \pi G \frac{\bar{\alpha}}{\bar{\nu}} \dot{\bar{\varphi}} \partial_c \delta\varphi^{\mathrm{GI}}$$

as the corrected time–space part of Einstein's equation,

$$\begin{aligned}
&\Delta(\bar{\alpha}^2\Psi) - 3\mathcal{H}(1+f)\left(\dot{\Psi} + \mathcal{H}\Phi(1+f)\right) \\
&\qquad = 4\pi G \frac{\bar{\alpha}}{\bar{\nu}} (1+f_3) \left(\dot{\bar{\varphi}} \delta\dot{\varphi}^{\mathrm{GI}} - \dot{\bar{\varphi}}^2 (1+f_1)\Phi + \bar{\nu} a^2 V_{,\varphi}(\bar{\varphi}) \delta\varphi^{\mathrm{GI}}\right)
\end{aligned}$$

as the corrected time–time part, and

$$\begin{aligned}
&\ddot{\Psi} + \mathcal{H}\left(2\dot{\Psi}\left(1 - \frac{a}{2\bar{\alpha}}\frac{\mathrm{d}\bar{\alpha}}{\mathrm{d}a}\right) + \dot{\Phi}(1+f)\right) \\
&\quad + \left(2\dot{\mathcal{H}} + \mathcal{H}^2\left(1 + \frac{a}{2}\frac{\mathrm{d}f}{\mathrm{d}a} - \frac{a}{2\bar{\alpha}}\frac{\mathrm{d}\bar{\alpha}}{\mathrm{d}a}\right)\right)\Phi(1+f) \\
&\qquad = 4\pi G \frac{\bar{\alpha}}{\bar{\nu}} \left(\dot{\bar{\varphi}} \delta\dot{\varphi}^{\mathrm{GI}} - a^2 \bar{\nu} V_{,\varphi}(\bar{\varphi}) \delta\varphi^{\mathrm{GI}}\right)
\end{aligned}$$

as the diagonal space–space part. All corrections f, f_1, f_3 and h below are determined from the basic inverse triad corrections $\bar{\alpha}$ (for the gravitational part of the constraint) and $\bar{\nu}$ (for the kinetic term of the matter part). Since these are background corrections, their form can easily be computed in isotropic models, suitably parameterized for all ambiguities and lattice refinement.

The off-diagonal space–space part also implies a non-trivial equation, with an unexpected consequence: while the classical analog would simply identify $\Phi = \Psi$, the corrected equation implies $\Phi = \Psi(1+h)$ with a quantum correction by $h \neq 0$. This may be interpreted as an effective anisotropic stress contribution, but it results from a correction to quantum gravity, not to matter.

As a second unexpected implication, we have non-conservation of power on large scales, as pointed out in [45, 49]. This may be important for inflationary structure formation, where the long evolution times while modes are outside the Hubble radius would make even a weakly changing size of the overall power significant. Both of these effects are difficult to see in gauge-fixed treatments, such as the longitudinal or uniform gauge, or in frame-fixed versions based on reduced phase-space quantizations. Since scalar cosmological perturbations have been formulated consistently only for inverse triad corrections, no version is yet able to evolve perturbations through a bounce for which holonomy corrections are required. (Implementing holonomy corrections only for the background leads to inconsistent evolution equations for inhomogeneities.)

In addition to inverse triad corrections, holonomies and quantum back-reaction must be implemented to obtain a full picture from loop quantum gravity. While consistent deformations are not yet known for the latter two corrections, [39] has formulated quantum back-reaction in a cosmological setting. In general, in quantum gravity this requires the inclusion of moments between all degrees of freedom of gravity and matter, including quantum correlations between them. By setting the quantum variables of gravity as well as its quantum correlations to zero, one obtains the effective equations of quantum field theory on a curved space-time as a limit. Including leading order corrections from the gravitational quantum variables provides quantum field theory on a quantum space-time. While such limiting cases can be realized explicitly at the effective level, much still remains to be done for a detailed analysis of specific properties.

11.5.3 Causality

Several examples illustrate the importance of having a consistent constraint algebra, rather than just any deformation as it would be allowed in gauge-fixed or frame-fixed treatments. We have already seen that some effects in cosmological perturbation equations can be seen only when neither gauge nor frame are fixed before the theory is quantized or effective constraints are derived. Another example is the realization of causality, as in [34] studied by comparing the propagation of gravitational waves to that of light.

With quantum gravity corrections, the gravitational as well as the Maxwell Hamiltonian change compared to the classical expressions, affecting evolution equations and their plane-wave solutions. The gravitational contribution to the Hamiltonian constraint for inverse triad corrections is

$$H_{\mathrm{G}} = \frac{1}{16\pi G}\int_{\Sigma} \mathrm{d}^3x \alpha(E^a_i) \frac{E^c_j E^d_k}{\sqrt{|\det E|}} \left(\varepsilon_i{}^{jk} F^i_{cd} - 2(1+\gamma^2) K^j_{[c} K^k_{d]}\right),$$

implying the linearized wave equation

$$\frac{1}{2}\left(\frac{1}{\alpha}\ddot{h}^i_a + 2\frac{\dot{a}}{a}\left(1 - \frac{2a\mathrm{d}\alpha/\mathrm{d}a}{\alpha}\right)\dot{h}^i_a - \alpha\nabla^2 h^i_a\right) = 8\pi G \Pi^i_a$$

for the tensor mode h^i_a on a cosmological background with scale factor a and source term Π^i_a. By a plane-wave ansatz, we derive the dispersion relation $\omega^2 = \alpha^2 k^2$ for gravitational waves. Since $\alpha > 1$ for perturbative corrections, there is a danger of gravitational waves being super-luminal.

With these corrections, gravitational waves are faster than light on a classical background. For a meaningful comparison, however, we should use the speed of gravitational waves in relation to the physical speed of light on the same background, which should receive quantum corrections, too. Here, the Hamiltonian is

$$H_{\rm EM} = \int_\Sigma {\rm d}^3x \left(\alpha_{\rm EM}(q_{cd}) \frac{2\pi}{\sqrt{q}} E^a E^b q_{ab} + \beta_{\rm EM}(q_{cd}) \frac{\sqrt{q}}{16\pi} F_{ab} F_{cd} q^{ac} q^{bd} \right)$$

with two correction functions $\alpha_{\rm EM}$ and $\beta_{\rm EM}$ kinematically independent of the gravitational correction α. From the wave equation

$$\partial_t \left(\alpha_{\rm EM}^{-1} \partial_t A_a \right) - \beta_{\rm EM} \nabla^2 A_a = 0$$

we obtain the dispersion relation $\omega^2 = \alpha_{\rm EM}\beta_{\rm EM}k^2$, which also is "super-luminal" compared to the classical speed of light.

Working out the requirements for anomaly-freedom with these two contributions to the Hamiltonian constraint, as done in [34], we find $\alpha^2 = \alpha_{\rm EM}\beta_{\rm EM}$ and the dispersion relations are equal. Physically, for comparisons of speeds on the same background, there is no super-luminal propagation. In a gauge-fixed or frame-fixed treatment, by comparison, one could have chosen the correction functions independently of each other since tensor modes of the gravitational field and the electric field make up independent physical observables. A gauge or frame-fixed treatment could easily produce corrected equations violating causality, but this is ruled out by a complete treatment.

11.6 Outlook: future dynamics

To probe a quantum theory of gravity or even arrive at predictions one must evaluate its dynamics in detail. For low-energy effects, leading corrections to classical equations must be derived. The best tool for systematic investigations in such situations is that of effective descriptions, providing the evolution of expectation values of observables in a physical state. At the same time, they can tell us much of the entire behavior of physical states.

Hamiltonian effective descriptions can be applied directly to canonical quantum gravity and cosmological models. In particular, typical implications such as the discreteness of spatial or space-time structures can then be probed, or first ensured to be consistent at all. Several examples in loop quantum cosmology have demonstrated that discreteness corrections can indeed be implemented consistently, leaving the theory covariant but deforming its local space-time symmetries. Future work must ensure that this is indeed possible for the full theory of loop quantum gravity and its effective constraints.

While isotropic solvable models of loop quantum cosmology suggest a role of bouncing cosmologies for potential scenarios, no consistent set of equations to evolve inhomogeneities through a bounce has been found. The only available options so far make use of gauge-fixings (or frame-fixings) before quantization, and thus miss crucial aspects of quantum space-time structures. Inhomogeneous cosmological scenarios remain uncertain, and with it follow-up issues such as the entropy problem.

What models investigated so far suggest, in many different versions, is that the classical algebra of space-time diffeomorphisms is deformed but not violated. It is then clear that quantum corrections cannot merely amount to higher curvature terms in an effective action, although such terms may appear, too. Instead, quantum structures of space-time must change by quantum effects. While space-time covariance is no longer realized in the standard sense, from the Hamiltonian perspective the effective theories remain completely consistent and covariant with an underlying first-class algebra of gauge generators. As indicated by the initial quote from [59], this sense of covariance, from a general perspective, is the appropriate one. Its consistent implementation, without fixing gauge or frame, can tell us a great deal about the fundamental structures of space and time.

References

[1] Artymowski, M., Lalak, Z. and Szulc, L. 2009. Loop quantum cosmology corrections to inflationary models. *JCAP*, **0901**, 004.

[2] Ashtekar, A. 1987. New Hamiltonian formulation of general relativity. *Phys. Rev. D*, **36**(6), 1587–602.

[3] Ashtekar, A. and Lewandowski, J. 1997. Quantum theory of geometry I: Area operators. *Class. Quantum Grav.*, **14**, A55–A82.

[4] Ashtekar, A. and Lewandowski, J. 1998. Quantum theory of geometry II: Volume operators. *Adv. Theor. Math. Phys.*, **1**, 388–429.

[5] Ashtekar, A. and Lewandowski, J. 2004. Background independent quantum gravity: A status report. *Class. Quantum Grav.*, **21**, R53–R152.

[6] Ashtekar, A., Lewandowski, J., Marolf, D., Mourão, J. and Thiemann, T. 1995. Quantization of diffeomorphism invariant theories of connections with local degrees of freedom. *J. Math. Phys.*, **36**(11), 6456–93.

[7] Ashtekar, A., Baez, J. C., Corichi, A. and Krasnov, K. 1998. Quantum geometry and black hole entropy. *Phys. Rev. Lett.*, **80**, 904–7.

[8] Ashtekar, A., Bojowald, M. and Lewandowski, J. 2003. Mathematical structure of loop quantum cosmology. *Adv. Theor. Math. Phys.*, **7**, 233–68.

[9] Ashtekar, A., Pawlowski, T. and Singh, P. 2006. Quantum nature of the Big Bang: An analytical and numerical investigation. *Phys. Rev. D*, **73**, 124038.

[10] Bahr, B. and Dittrich, B. 2009a. Breaking and restoring of diffeomorphism symmetry in discrete gravity.

[11] Bahr, B. and Dittrich, B. 2009b. Improved and perfect actions in discrete gravity.

[12] Banerjee, K. and Date, G. 2005. Discreteness corrections to the effective Hamiltonian of isotropic loop quantum cosmology. *Class. Quant. Grav.*, **22**, 2017–33.

[13] Barbero, J. F. 1995. Real Ashtekar variables for Lorentzian signature space-times. *Phys. Rev. D*, **51**(10), 5507–10.

[14] Bardeen, J. M. 1980. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, **22**, 1882–905.
[15] Barrau, A. and Grain, J. 2009. Cosmological footprint of loop quantum gravity. *Phys. Rev. Lett.*, **102**, 081301.
[16] Bentivegna, E. and Pawlowski, T. 2008. Anti-deSitter universe dynamics in LQC. *Phys. Rev. D*, **77**, 124025.
[17] Bergmann, P. G. 1961. Observables in general relativity. *Rev. Mod. Phys.*, **33**, 510–14.
[18] Bojowald, M. 2001a. Absence of a singularity in loop quantum cosmology. *Phys. Rev. Lett.*, **86**, 5227–30.
[19] Bojowald, M. 2001b. Inverse scale factor in isotropic quantum geometry. *Phys. Rev. D*, **64**, 084018.
[20] Bojowald, M. 2001c. Loop quantum cosmology IV: Discrete time evolution. *Class. Quantum Grav.*, **18**, 1071–88.
[21] Bojowald, M. 2002a. Isotropic loop quantum cosmology. *Class. Quantum Grav.*, **19**, 2717–41.
[22] Bojowald, M. 2002b. Quantization ambiguities in isotropic quantum geometry. *Class. Quantum Grav.*, **19**, 5113–30.
[23] Bojowald, M. 2004. Spherically symmetric quantum geometry: states and basic operators. *Class. Quantum Grav.*, **21**, 3733–53.
[24] Bojowald, M. 2006. Loop quantum cosmology and inhomogeneities. *Gen. Rel. Grav.*, **38**, 1771–95.
[25] Bojowald, M. 2007a. Dynamical coherent states and physical solutions of quantum cosmological bounces. *Phys. Rev. D*, **75**, 123512.
[26] Bojowald, M. 2007b. Large scale effective theory for cosmological bounces. *Phys. Rev. D*, **75**, 081301(R).
[27] Bojowald, M. 2007c. What happened before the big bang? *Nature Physics*, **3**(8), 523–5.
[28] Bojowald, M. 2008a. The dark side of a patchwork universe. *Gen. Rel. Grav.*, **40**, 639–60.
[29] Bojowald, M. 2008b. How quantum is the big bang? *Phys. Rev. Lett.*, **100**, 221301.
[30] Bojowald, M. 2008c. Loop quantum cosmology. *Living Rev. Relativity*, **11**, 4. http://www.livingreviews.org/lrr-2008-4.
[31] Bojowald, M. 2008d. Quantum nature of cosmological bounces. *Gen. Rel. Grav.*, **40**, 2659–83.
[32] Bojowald, M. and Das, R. 2008. Fermions in loop quantum cosmology and the role of parity. *Class. Quantum Grav.*, **25**, 195006.
[33] Bojowald, M. and Hossain, G. 2007. Cosmological vector modes and quantum gravity effects. *Class. Quantum Grav.*, **24**, 4801–16.
[34] Bojowald, M. and Hossain, G. 2008. Quantum gravity corrections to gravitational wave dispersion. *Phys. Rev. D*, **77**, 023508.
[35] Bojowald, M. and Kagan, M. 2006. Singularities in isotropic non-minimal scalar field models. *Class. Quantum Grav.*, **23**, 4983–90.
[36] Bojowald, M. and Kastrup, H. A. 2000. Symmetry reduction for quantized diffeomorphism invariant theories of connections. *Class. Quantum Grav.*, **17**, 3009–43.
[37] Bojowald, M. and Reyes, J. D. 2009. Dilaton gravity, Poisson sigma models and loop quantum gravity. *Class. Quantum Grav.*, **26**, 035018.
[38] Bojowald, M. and Skirzewski, A. 2006. Effective equations of motion for quantum systems. *Rev. Math. Phys.*, **18**, 713–45.
[39] Bojowald, M. and Skirzewski, A. 2008. Effective theory for the cosmological generation of structure. *Adv. Sci. Lett.*, **1**, 92–8.

[40] Bojowald, M. and Strobl, T. 2003. Poisson geometry in constrained systems. *Rev. Math. Phys.*, **15**, 663–703.
[41] Bojowald, M. and Tavakol, R. 2008. Recollapsing quantum cosmologies and the question of entropy. *Phys. Rev. D*, **78**, 023515.
[42] Bojowald, M. and Tsobanjan, A. 2009. Effective constraints for relativistic quantum systems. *Phys. Rev. D*, to appear.
[43] Bojowald, M., Hernández, H. H. and Morales-Técotl, H. A. 2006. Perturbative degrees of freedom in loop quantum gravity: Anisotropies. *Class. Quantum Grav.*, **23**, 3491–516.
[44] Bojowald, M., Hernández, H. and Skirzewski, A. 2007a. Effective equations for isotropic quantum cosmology including matter. *Phys. Rev. D*, **76**, 063511.
[45] Bojowald, M., Hernández, H., Kagan, M., Singh, P. and Skirzewski, A. 2007b. Formation and evolution of structure in loop cosmology. *Phys. Rev. Lett.*, **98**, 031301.
[46] Bojowald, M., Cartin, D. and Khanna, G. 2007c. Lattice refining loop quantum cosmology, anisotropic models and stability. *Phys. Rev. D*, **76**, 064018.
[47] Bojowald, M., Hossain, G., Kagan, M. and Shankaranarayanan, S. 2008. Anomaly freedom in perturbative loop quantum gravity. *Phys. Rev. D*, **78**, 063547.
[48] Bojowald, M., Sandhöfer, B., Skirzewski, A. and Tsobanjan, A. 2009a. Effective constraints for quantum systems. *Rev. Math. Phys.*, **21**, 111–54.
[49] Bojowald, M., Hossain, G., Kagan, M. and Shankaranarayanan, S. 2009b. Gauge invariant cosmological perturbation equations with corrections from loop quantum gravity. *Phys. Rev. D*, **79**, 043505.
[50] Bojowald, M., Reyes, J. D. and Tibrewala, R. 2009c. Non-marginal LTB-like models with inverse triad corrections from loop quantum gravity. *Phys. Rev. D*, **80**, 084002.
[51] Bojowald, M. 2009. Consistent loop quantum cosmology. *Class. Quantum Grav.*, **26**, 075020.
[52] Bruni, M., Dunsby, P. K. S. and Ellis, G. F. R. 1992. Cosmological perturbations and the physical meaning of gauge invariant variables. *Astrophys. J.*, **395**, 34–53.
[53] Brunnemann, J. and Fleischhack, C. 2007. On the configuration spaces of homogeneous loop quantum cosmology and loop quantum gravity.
[54] Cametti, F., Jona-Lasinio, G., Presilla, C. and Toninelli, F. 2000. Comparison between quantum and classical dynamics in the effective action formalism. Pages 431–48 of: *Proceedings of the International School of Physics "Enrico Fermi", Course CXLIII*. Amsterdam: IOS Press.
[55] Campiglia, M., Di Bartolo, C., Gambini, R. and Pullin, J. 2006. Uniform discretizations: a new approach for the quantization of totally constrained systems. *Phys. Rev. D*, **74**, 124012.
[56] Campiglia, M., Gambini, R. and Pullin, J. 2007. Loop quantization of spherically symmetric midi-superspaces. *Class. Quantum Grav.*, **24**, 3649.
[57] Copeland, E. J., Mulryne, D. J., Nunes, N. J. and Shaeri, M. 2009. The gravitational wave background from super-inflation in Loop Quantum Cosmology. *Phys. Rev. D*, **79**, 023508.
[58] Deruelle, N., Sasaki, M., Sendouda, Y. and Yamauchi, D. 2009. Hamiltonian formulation of f(Riemann) theories of gravity.
[59] Dirac, P. A. M. 1958. The theory of gravitation in Hamiltonian form. *Proc. Roy. Soc. A*, **246**, 333–43.
[60] Dittrich, B. 2006. Partial and complete observables for canonical general relativity. *Class. Quant. Grav.*, **23**, 6155–84.
[61] Dittrich, B. 2007. Partial and complete observables for Hamiltonian constrained systems. *Gen. Rel. Grav.*, **39**, 1891–927.

[62] Domagala, M. and Lewandowski, J. 2004. Black hole entropy from quantum geometry. *Class. Quantum Grav.*, **21**, 5233–43.

[63] Ellis, G. F. R. and Bruni, M. 1989. Covariant and gauge invariant approach to cosmological density fluctuations. *Phys. Rev. D*, **40**, 1804–18.

[64] Ellis, G. F. R. and Maartens, R. 2004. The emergent universe: inflationary cosmology with no singularity. *Class. Quant. Grav.*, **21**, 223–32.

[65] Ellis, G. F. R., Murugan, J. and Tsagas, C. G. 2004. The emergent universe: An explicit construction. *Class. Quant. Grav.*, **21**, 233–50.

[66] Fewster, C. and Sahlmann, H. 2008. Phase space quantization and loop quantum cosmology: A Wigner function for the Bohr-compactified real line. *Class. Quantum Grav.*, **25**, 225015.

[67] Fleischhack, C. 2009. Representations of the Weyl algebra in quantum geometry. *Commun. Math. Phys.*, **285**, 67–140.

[68] Giesel, K., Hofmann, S., Thiemann, T. and Winkler, O. 2007a. Manifestly gauge-invariant general relativistic perturbation theory: I. Foundations.

[69] Giesel, K., Hofmann, S., Thiemann, T. and Winkler, O. 2007b. Manifestly gauge-invariant general relativistic perturbation theory: II. FRW Background and first order.

[70] Giesel, K., Tambornino, J. and Thiemann, T. 2009. LTB spacetimes in terms of Dirac observables.

[71] Grain, J., Cailleteau, T., Barrau, A. and Gorecki, A. 2009a. Fully LQC-corrected propagation of gravitational waves during slow-roll inflation.

[72] Grain, J., Barrau, A. and Gorecki, A. 2009b. Inverse volume corrections from loop quantum gravity and the primordial tensor power spectrum in slow-roll inflation. *Phys. Rev. D*, **79**, 084015.

[73] Husain, V. and Winkler, O. 2004. On singularity resolution in quantum gravity. *Phys. Rev. D*, **69**, 084016.

[74] Immirzi, G. 1997. Real and complex connections for canonical gravity. *Class. Quantum Grav.*, **14**, L177–L181.

[75] Jacobson, T. 2000. Trans-Planckian redshifts and the substance of the space-time river.

[76] Kaul, R. K. and Majumdar, P. 1998. Quantum black hole entropy. *Phys. Lett. B*, **439**, 267–70.

[77] Kibble, T. W. B. 1979. Geometrization of quantum mechanics. *Commun. Math. Phys.*, **65**, 189–201.

[78] Laddha, A. 2007. Polymer quantization of CGHS model – I. *Class. Quant. Grav.*, **24**, 4969–88.

[79] Laddha, A. and Varadarajan, M. 2008. Polymer parametrised field theory. *Phys. Rev. D*, **78**, 044008.

[80] Lewandowski, J., Okolów, A., Sahlmann, H. and Thiemann, T. 2006. Uniqueness of diffeomorphism invariant states on holonomy-flux algebras. *Commun. Math. Phys.*, **267**, 703–33.

[81] Martin-Benito, M., Garay, L. J. and Mena Marugán, G. A. 2008. Hybrid quantum Gowdy cosmology: Combining loop and Fock quantizations. *Phys. Rev. D*, **78**, 083516.

[82] Meissner, K. A. 2004. Black hole entropy in loop quantum gravity. *Class. Quantum Grav.*, **21**, 5245–51.

[83] Mielczarek, J. 2008. Gravitational waves from the Big Bounce. *JCAP*, **0811**, 011.

[84] Mielczarek, J. 2009. The observational implications of loop quantum cosmology.

[85] Nelson, W. and Sakellariadou, M. 2007a. Lattice refining loop quantum cosmology and inflation. *Phys. Rev. D*, **76**, 044015.

[86] Nelson, W. and Sakellariadou, M. 2007b. Lattice refining LQC and the matter Hamiltonian. *Phys. Rev. D*, **76**, 104003.
[87] Nelson, W. and Sakellariadou, M. 2008. Numerical techniques for solving the quantum constraint equation of generic lattice-refined models in loop quantum cosmology. *Phys. Rev. D*, **78**, 024030.
[88] Puchta, J. 2009. Ph.D. thesis, University of Warsaw.
[89] Reyes, J. D. 2009. *Spherically Symmetric Loop Quantum Gravity: Connections to 2-Dimensional Models and Applications to Gravitational Collapse*. Ph.D. thesis, The Pennsylvania State University.
[90] Rovelli, C. 1991a. Quantum reference systems. *Class. Quantum Grav.*, **8**, 317–32.
[91] Rovelli, C. 1991b. What is observable in classical and quantum gravity? *Class. Quantum Grav.*, **8**, 297–316.
[92] Rovelli, C. 2004. *Quantum Gravity*. Cambridge, UK: Cambridge University Press.
[93] Rovelli, C. and Smolin, L. 1990. Loop space representation of quantum general relativity. *Nucl. Phys. B*, **331**, 80–152.
[94] Rovelli, C. and Smolin, L. 1994. The physical Hamiltonian in nonperturbative quantum gravity. *Phys. Rev. Lett.*, **72**, 446–9.
[95] Rovelli, C. and Smolin, L. 1995. Discreteness of area and volume in quantum gravity. *Nucl. Phys. B*, **442**, 593–619. Erratum: *Nucl. Phys. B* **456** (1995) 753.
[96] Rovelli, C. and Vidotto, F. 2008. Stepping out of homogeneity in loop quantum cosmology. *Class. Quantum Grav.*, **25**, 225024.
[97] Sabharwal, S. and Khanna, G. 2008. Numerical solutions to lattice-refined models in loop quantum cosmology. *Class. Quantum Grav.*, **25**, 085009.
[98] Sahlmann, H. 2009. This volume.
[99] Shimano, M. and Harada, T. 2009. Observational constraints of a power spectrum from super-inflation in loop quantum cosmology.
[100] Singh, P. 2006. Loop cosmological dynamics and dualities with Randall–Sundrum braneworlds. *Phys. Rev. D*, **73**, 063508.
[101] Singh, P. and Vandersloot, K. 2005. Semi-classical states, effective dynamics and classical emergence in loop quantum cosmology. *Phys. Rev. D*, **72**, 084004.
[102] Taveras, V. 2008. Corrections to the Friedmann equations from LQG for a universe with a free scalar field. *Phys. Rev. D*, **78**, 064072.
[103] Thiemann, T. 1998a. Quantum spin dynamics (QSD). *Class. Quantum Grav.*, **15**, 839–73.
[104] Thiemann, T. 1998b. QSD V: Quantum gravity as the natural regulator of matter quantum field theories. *Class. Quantum Grav.*, **15**, 1281–314.
[105] Thiemann, T. 2007. *Introduction to Modern Canonical Quantum General Relativity*. Cambridge, UK: Cambridge University Press.
[106] Unruh, W. 1997. *Time, Gravity, and Quantum Mechanics*. Cambridge, UK: Cambridge University Press, pp. 23–94.
[107] Weiss, N. 1985. Constraints on Hamiltonian lattice formulations of field theories in an expanding universe. *Phys. Rev. D*, **32**, 3228–32.

12

The microscopic dynamics of quantum space as a group field theory

DANIELE ORITI

We provide a rather extended introduction to the group field theory approach to quantum gravity, and the main ideas behind it. We present in some detail the GFT quantization of 3D Riemannian gravity, and discuss briefly the current status of the 4-dimensional extensions of this construction. We also briefly report on some recent results, concerning both the mathematical definition of GFT models as bona fide field theories, and avenues towards extracting testable physics from them.

12.1 Introduction

The field of non-perturbative and background-independent quantum gravity has progressed considerably over the past few decades [78]. New research directions are being developed, new important developments are taking place in existing approaches, and some of these approaches are converging to one another. As a result, ideas and tools from one become relevant to another, and trigger further progress. The group field theory (GFT) formalism [39, 77, 79] nicely captures this convergence of approaches and ideas. It is a generalization of the much studied matrix models for 2D quantum gravity and string theory [28, 53]. At the same time, it generalizes it, as we are going to explain, by incorporating the insights coming from canonical loop quantum gravity and its covariant spin foam formulation of the dynamics, and so it became an important part of this approach to the quantization of 4D gravity [72, 74, 81, 85]. Furthermore, it is a point of convergence of the same loop quantum gravity approach and of simplicial quantum gravity approaches, like quantum Regge calculus [93] and dynamical triangulations [3, 79], in that the covariant dynamics of the first takes the form, as we are going to see,

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

of simplicial path integrals. More recently, tools and ideas from non-commutative geometry have been introduced as well in the formalism, and this has facilitated the attempts to make tentative contact with effective models and with quantum gravity phenomenology.

The aim of this chapter is to explain the general idea behind the GFT formalism and its roots, discuss its relation with other current approaches to quantum gravity, detail to some extent the construction and features of GFT models in three and four dimensions, and finally report briefly on some recent results, concerning both an improved mathematical understanding of it and results with possible bearing on phenomenology. The models we will discuss in some more detail are in Euclidean signature, but the whole construction can straightforwardly be performed in the Lorentzian setting as well, and many results are known about the corresponding Lorentzian GFT models.

12.1.1 A general definition

In very general terms, group field theories are an attempt to define quantum gravity in terms of *combinatorially non-local quantum field theories on group manifolds or on the corresponding Lie algebras*, related to the Lorentz or rotation group. The formalism itself is mostly characterized by the combinatorial non-locality that we are going to specify in the following, and the choice of group manifolds as domain of definition of the field, and in particular of the Lorentz/rotation groups, is dictated by the wish to model quantum gravity. Other choices could be devised easily, for different purposes (e.g. describing matter or gauge fields, capturing topological structures, etc.). We stick to quantum gravity models in this chapter. Before introducing the formalism or specific models, let us motivate in some detail these basic ingredients: a quantum field theory framework, the use of group structures, the combinatorial non-locality.

12.1.2 A quantum field theory for quantum gravity?

It is actually easy to see why we may want to use a quantum field theory formalism (QFT) also for quantum gravity. Quantum field theory is the best formalism we have for describing physics at both microscopic and mesoscopic scales, both in high energy particle physics and in condensed matter physics, for both elementary systems and many-particle ones. And actually, even at large and very large scales, it is still field theory that we use, only we are most often able to neglect quantum aspects (e.g. general relativity itself).

So, probably, a more relevant question is: can we still hope for a formulation of quantum gravity, for description of the microscopic structure of quantum space, in terms of a quantum field theory? The rationale behind such a question is the following. We only know how to define quantum field theories on given background manifolds, and we have full control (including renormalization, etc.) only over quantum field theories on flat spaces. Moreover, we have tried already to apply this formalism to gravity, formulating it as a quantum field theory of massless spin-2 particles propagating on a flat space, the gravitons, thought of as the carriers of the gravitational interaction, coupled universally to other matter and gauge fields. It has been the first approach to quantum gravity ever developed, by the great scientist of the past century (Rovelli [84]). We know it does not work. The field theory defined by this approximation is not renormalizable. Quantum gravity, beyond the effective field theory level, is not such a quantum field theory.

This historically well-founded objection, however, is not a no-go theorem, of course. In particular, the non-renormalizability result does not rule out the use of a quantum field theory formalism as such. What it rules out is the specific idea of a field theory of gravitons as a fundamental definition of the dynamics of quantum space. To the eyes of many, it rules out also the idea that the requirement of manifest background independence with respect to spacetime is a dispensable one, supporting instead the belief that this should be the defining property of any sound quantum gravity formalism [51, 86].

The more serious objection to the idea of using a quantum field theory formalism for quantum gravity, in fact, is that a good theory of quantum gravity should be background-independent, because it should explain the origin and properties *of spacetime* itself, of its geometry and, maybe, its topology. But, as said, we know how to formulate quantum field theories only on given backgrounds.

Again, this does not rule out the use of the QFT formalism. It means, however, that if an (almost) ordinary QFT it should be, quantum gravity can only be a QFT on some auxiliary, internal or "higher-level" space.

So we can then look at general relativity (GR) itself and try to identify some background (non-dynamical) structures that could provide such space. After all, GR *is* a classical field theory, and it is background-independent with respect to the geometry (not the topology) of spacetime.

12.1.3 Background structures in classical and quantum GR

The first background structure that comes to mind is the spacetime dimension. We have of course overwhelming experimental evidence for a 4-dimensional spacetime. But it is also true that we do not have a clear enough understanding of *why* this dimensionality should hold true at high energy, microscopic length scales or when all quantum effects of space dynamics are taken into account. So it makes sense to look for alternatives, i.e. for a formalism in which the spacetime dimension is dynamical. Group field theories (as loop quantum gravity or simplicial quantum gravity and tensor models) fix the kinematical dimension of space at the very beginning, at least in the present formulation. However, on the one hand they do not suggest any obstruction to dimensional generalizations of the formalism. On the other hand, this kinematical choice ensures that the dynamic dimension of spacetime in some effective continuum and classical description will match the kinematical one. The example of dynamical triangulations [3], in fact, shows that this matching is far from trivial.

Another background structure is the internal, local symmetry group of the theory, i.e. the Lorentz group, which provides the local invariance under change of reference (tetrad) frame, and that is at the heart of the equivalence principle. It is a sort of "background internal space" of the theory. This gives the primary conceptual motivation for using the Lorentz group (and related) in GFT. At the same time, as we are going to discuss in the following, this choice allows us to incorporate in the GFT formalism what we have learned from canonical loop quantum gravity, as well as many of its results [85, 90]. In fact, the role of the Lorentz group (and of its $SU(2)$ rotation subgroup) is brought to the forefront in the LQG formalism. LQG is based on the classical reformulation of GR as a background-independent (and diffeomorphism-invariant) theory of a Lorentz connection. Upon quantization, it gives a space of states which is an L^2 space of generalized $SU(2)$ connections, obtained as the projective limit of the space of L^2 cylindrical functions of finite numbers of $SU(2)$ group elements, representing parallel transports of the same connection along elementary paths in space. Thus spacetime geometry is encoded, ultimately, in these group-theoretic structures.

A background structure of GR is, in fact, also its configuration space, seen from the Hamiltonian perspective, and regardless of the specific variables used to parametrize it: the space $\mathcal{S}$ of (spatial) geometries on a given (spatial) topology, coined "superspace" by Wheeler. In the ADM variables, this is a metric space in its own right [55] and could be considered indeed a sort of "background meta-space: a space of spaces." Let us sketch briefly how this background structure enters the quantization of the classical theory, at least at the formal level, in both canonical and covariant approaches. Loop quantum gravity, spin foam models, and simplicial quantum gravity reformulate and make more precise and successful, to different degrees, these "historic" ideas.

The canonical approach starts with a globally hyperbolic spacetime, with topology $\Sigma \times \mathbb{R}$. For simplicity, Σ is usually chosen to be compact and simply connected, with the topology of the 3-sphere S^3. The wave functions of canonical geometrodynamics are functionals on the space of 3-metrics on the 3-sphere, $\Psi(h_{ij})$, and kinematical observables are functionals of the phase-space variables, themselves built from the conjugate 3-metric h_{ij} and extrinsic curvature K_{ij}, turned into (differential) operators acting on the wave functions. Gravity being (classically) a totally constrained system, the dynamics is imposed by identifying the space of states (and associated inner product), i.e. the space of functionals on the space of 3-geometries (metrics up to spatial diffeomorphisms), which satisfy also the Hamiltonian constraint, and thus the Wheeler–deWitt equation $\mathcal{H}\Psi(h_{ij}) = 0$. This, together with the identification of physical observables, defines the theory from a canonical perspective. A covariant definition of the dynamics can instead be looked for in sum-over-histories framework. Given the same (trivial) spacetime topology, and the same choice of spatial topology, consider all the possible geometries (spacetime metrics up to diffeomorphisms) that are compatible with it. Transition amplitudes (defining either the physical inner product of the canonical theory or "causal" transition amplitudes, and thus Green functions for the canonical Hamiltonian constraint operator [89]), for given boundary configurations of the field, i.e. possible 3-geometries on the 3-sphere: h and h', would be given by a sum over spacetime geometries like:

$$Z_{QG}\left(h, h'\right) = \int_{g(\mathcal{M}|h,h')} \mathcal{D}g\, e^{i\, S_{GR}(g,\mathcal{M})}, \tag{12.1}$$

i.e. by summing over all 4-geometries inducing the given 3-geometries on the boundary, with the amplitude possibly modified by boundary terms if needed. The expression above is obviously purely formal, for a variety of well-known reasons. In any case, it looks like a prototype of a background-independent quantization of spacetime geometry, for given spatial and spacetime topology, and given space $\mathcal{S}$ of possible 3-geometries. Also, the physical interpretation of the above quantities presents several challenges, given that the formalism seems to be bound to a cosmological setting, where our usual interpretation of quantum mechanics is rather dubious. A good point about group field theory, and about LQG, spin foams and simplicial gravity, is that they seem to provide a more rigorous definition of the above formulas, which is also local in a sense to be clarified below.

12.1.4 Making topology dynamical: the idea of third quantization

In spite of the difficulties in making sense of a path integral quantization of gravity on a fixed spacetime, one may wish also to treat topology as a dynamical variable in the theory. One would therefore try to implement a sort of "sum over topologies" alongside a sum over geometries, thus extending this last sum to run over *all* possible spacetime geometries and not only those that can live on a given topology. This has consequences for the type of geometries one can consider, in the Lorentzian case, given that a non-trivial spacetime topology implies spatial topology change [34] and this in turn forces the metric to allow either for closed timelike loops or for geometries which are degenerate (at least) at isolated points. This argument was made stronger by Horowitz [59], concluding that if degenerate metrics are included in the (quantum) theory, then topology change is not only possible but unavoidable and non-trivial topologies therefore must be included in the quantum theory. There are several other results that suggest the need for topology change in quantum gravity, including work on topological geons [33], in string theory [5], and on wormholes as a possible explanatory mechanism for the small value of the cosmological constant [7], and the possibility has been raised that *all* constants of nature can be seen as computable dynamical vacuum expectation values in a theory in which topology change is allowed [25].

This last idea, together with the analogy with string perturbation theory and the aim to solve some problems of the canonical formulation of quantum gravity, prompted the proposal of a "third quantization" formalism for quantum gravity [52, 65, 70]. The general idea is to define a (scalar) field theory on superspace $\mathcal{S}$ for a given choice of basic spatial manifold topology, e.g. the 3-sphere, essentially turning the wave function of the canonical theory into an operator: $\phi(h)$, whose dynamics is defined by an action of the type

$$S(\phi) = \int_{\mathcal{S}} \mathcal{D}h\, \phi(h) \Delta \phi(h) + \lambda \int \mathcal{V}(\phi(h)) \tag{12.2}$$

with Δ being the Wheeler–DeWitt differential operator of canonical gravity, here defining the kinetic term (free propagation) of the theory, while $\mathcal{V}(\phi)$ is a generic, e.g. cubic, and generically non-local (in superspace) interaction term for the field, governing the topology-changing processes. Notice that because of the choice of basic spatial topology needed to define the third quantized field, the topology-changing processes described here are those turning X copies of the 3-sphere into Y copies of the same.

The quantum theory is "defined" by the partition function $Z = \int \mathcal{D}\phi e^{-S(\phi)}$, in its perturbative expansion in "Feynman diagrams" (Figure 12.1).

The Feynman amplitudes will be given by the quantum gravity path integral (sum over geometries) for each spacetime topology (identified with a particular

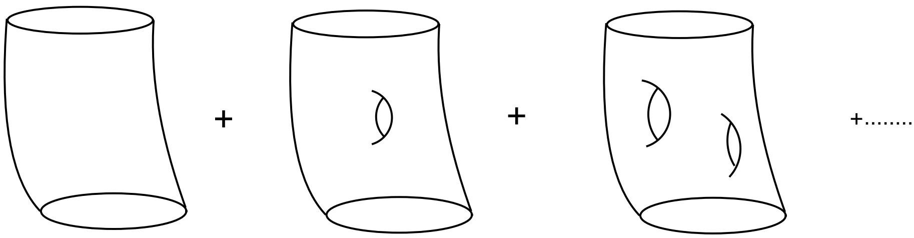


Figure 12.1 The perturbative expansion of the third quantized field theory in interaction processes for universes.

interaction process of universes). The one for trivial topology representing a sort of one-particle propagator, thus a Green function for the Wheeler–DeWitt equation. Other features of this (very) formal setting are: (1) the full classical equations of motion for the scalar field on superspace will be a non-linear extension of the Wheeler–DeWitt equation of canonical gravity, due to the interaction term in the action, i.e. the inclusion of topology change; (2) the perturbative third quantized vacuum of the theory will be the "no spacetime" state, and not any state with a semi-classical geometric interpretation in terms of a smooth geometry, say a Minkowski state. These features are shared by the group field theory approach, as we will see.

Notice that in this formalism for spacetime topology change, the spatial topology of each single-universe sector remains fixed, and the superspace $\mathcal{S}$ itself remains a background structure for the "third quantized" field theory.

Notice also that, if one was to attempt to reproduce this type of setting in terms of the variables used in LQG, two background structures of classical GR, i.e. the Lorentz group and the superspace $\mathcal{S}$, would be somehow unified, as superspace would have to be identified with the space of (generalized) Lorentz (or $\mathrm{SU}(2)$) connections on some given spatial topology. Again, something of this sort happens in group field theory, which can be seen as a sort of discrete non-local field theory on the space of geometries for *building blocks* of space, in turn given by group or Lie algebra variables. Before getting to the details of the GFT formalism, however, we want to motivate further the use of discrete structures and the associated non-locality.

12.1.5 A finitary substitute for continuum spacetime?

However good the idea of a path integral for gravity and its extension to a third quantized formalism may be, there has been no definite success in realizing them rigorously.

One is tempted to identify the main reason for the difficulties encountered in the use of a *continuum* for describing spacetime, both at the topological and at the geometrical level. One can advocate the use of *discrete structures* as a way to regularize and make computable the above expressions, to provide a more rigorous and fundamental definition of the theory, with the continuum description emerging only in a continuum *approximation* of the corresponding discrete quantities, like hydrodynamics for large ensembles of many particles. This was in fact among the motivations for discrete approaches to quantum gravity as matrix models, or dynamical triangulations or quantum Regge calculus. Indeed, various arguments have been put forward to support the point of view that discrete structures provide a more *fundamental* description of spacetime. One possibility, suggested by various approaches to quantum gravity such as string theory or loop quantum gravity, is the existence of a fundamental length scale that sets a lower bound for distances and thus makes the notion of a continuum lose its physical meaning. Also, one can argue on both philosophical and mathematical grounds [61] that the very notion of "point" can correspond at most to an idealization of the nature of spacetime (as implemented mathematically in non-commutative models of quantum gravity [4, 69]). Spacetime points are indeed to be replaced, from this point of view, by small but finite regions and a more fundamental model of spacetime should take these local regions as basic building blocks. Also, the results of black hole thermodynamics seem to suggest that there should be a finite number of fundamental spacetime degrees of freedom associated with any region of spacetime, the apparent continuum being the result of the microscopic (Planckian) nature of them [88]. In other words, a finitary topological space [87] would constitute a better model of spacetime than a smooth manifold. These arguments also favor a simplicial description of spacetime, with the simplices being a finitary substitute of points. And these same arguments are reinforced by the results of LQG whose kinematical states are labeled by graphs [90].

Here is where GFTs provide a discrete or finitary implementation of the third quantization idea, which is also a more local one. In GFT the spatial manifold is actually to be thought of as a collection of (glued) building blocks, akin to a many-particle state, and the field theory should be defined on the space of possible geometries of each such building block. Spatial topology is also allowed to change arbitrarily, in principle, if the building blocks are allowed to combine arbitrarily by the theory. These building blocks can be depicted as either fundamental simplices or (pieces of) finite graphs, as we will see, and the space of geometries of such discrete structures is then necessarily finite-dimensional, and it will be characterized, as said, by either group or Lie algebra elements.

So the GFT formalism incorporates insights from other approaches (loop quantum gravity and simplicial gravity) also in answering a second natural question that comes to mind when suggesting a quantum field theory formalism for the microstructure of space: a QFT of which fundamental quanta? Again, we know that these cannot be gravitons. They have to be quanta *of space* itself, excitations around a vacuum that corresponds to the *absence of space*.

To introduce how their dynamics is identified, and in the process motivate the peculiar *non-locality* of GFT interactions, we take a short detour and discuss first briefly their lower-dimensional predecessors: matrix models.

12.1.6 A combinatorial non-locality: from point particles to extended combinatorial structures

Consider a point particle in 0+1 dimensions, with action $S(X) = \frac{1}{2}X^2 + \frac{\lambda}{3}X^3$. This action defines a trivial dynamics (for a trivial system), of course. What interests us here, however, is the combinatorial structure of its "Feynman diagrams," i.e. the graphs that can be used as a convenient book-keeping tool in computing the corresponding partition function $Z = \int dX\, e^{-S_\lambda(X)}$ perturbatively in λ. These are simple 3-valent (because of the order of the "interaction") graphs.

The fact that the Feynman diagrams of the theory are simple graphs like in Figure 12.2 follows from (1) the *point-like* nature of the particle, and (2) the *locality* of the corresponding interaction, encoded in the identification of X variables in the interaction part of the action. A less trivial system would be given by a relativistic particle, for example, or, better, a system of such interacting particles. Notice also that the relativistic particle is often taken as a sort of general relativity in $0+1$ dimensions, and while this analogy has strong limitations (like all analogies), it is indeed very useful to grasp an intuitive understanding of several issues that show up in the (quantum) gravity context [89]. The combinatorial structure of the Feynman diagrams, now weighted by non-trivial amplitudes (convolutions of Feynman propagators for each interacting particle with identified initial/end points), will be the same as long as we do not allow for non-local interactions.

The same structure of diagrams is maintained, because the local nature of the interaction and the point-like nature of the corresponding quanta are maintained, also when moving to a field theory setting, i.e. going from the above particle dynamics to the corresponding field theory (still dynamically rather trivial), thus allowing for the creation and annihilation of particles. Infinite numbers of degrees of freedom are now possible, depending on the specific dynamics chosen, because of the arbitrary number of particles that can be involved in any interaction process. Assume for simplicity that the dynamics is governed by the (trivial)

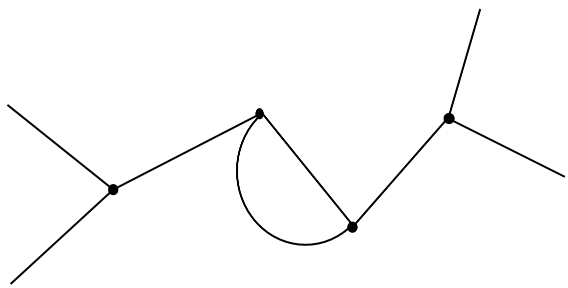


Figure 12.2 A Feynman graph for a point particle and the corresponding field theory.

action: $S(\phi) = \frac{1}{2}\int dx\, \phi(x)^2 + \frac{\lambda}{3}\int dx\, \phi(x)^3$. Now we have integrations over the position variables labeling the vertices, or, in momentum space, lines are *labeled* by momenta that sum to zero at vertices, and that are integrated over, reflecting a (potential) infinity of degrees of freedom in the theory. But still, the combinatorics of the diagrams is the same as in the corresponding particle case.

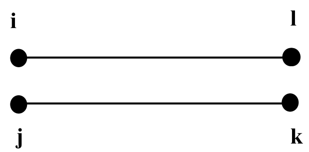
Matrix models and discrete and continuum 2D gravity

Now we move up in combinatorial dimension. Instead of point particles, let us consider 1-dimensional objects, that could be represented graphically by a line, with two end points. We label these two end points with two indices i, j, and we represent these fundamental objects of our theory by $N \times N$ *matrices* M_{ij} (with $i, j = 1, \ldots, N$) [28, 53], i.e. arrays of real or complex numbers replacing the "point" variables X. For simplicity, assume these matrices to be hermitian.

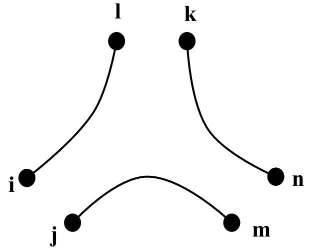
We want to move up in dimension also in the corresponding Feynman diagrams, i.e. we want to have diagrams that correspond to 2-dimensional structures, instead of 1-dimensional graphs. In order to do so, we have to drop the assumption of *locality*. We define a simple action for M, given by

$$
\begin{aligned}
S(M) &= \frac{1}{2} tr M^2 - \frac{g}{\sqrt{N}} tr M^3 = \frac{1}{2} M^i{}_j M^j{}_i - \frac{g}{\sqrt{N}} M^i{}_j M^j{}_k M^k{}_i \\
&= \frac{1}{2} M^i{}_j K_{jkli} M^k{}_l - \frac{g}{\sqrt{N}} M^i{}_j M^m{}_n M^k{}_l V_{jmknli} \\
&\quad \text{with} \quad K_{jkli} = \delta^j{}_k \delta^l{}_i, \quad V_{jmknli} = \delta^j{}_m \delta^n{}_k \delta^l{}_i, \quad \left(K^{-1}\right)_{jkli} = K_{jkli}.
\end{aligned}
$$

We have also identified in the above formula the propagator and vertex term that will give the building blocks of the corresponding Feynman amplitudes. These building blocks can be represented graphically as follows:

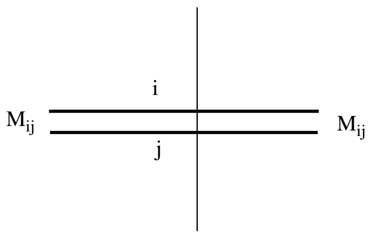


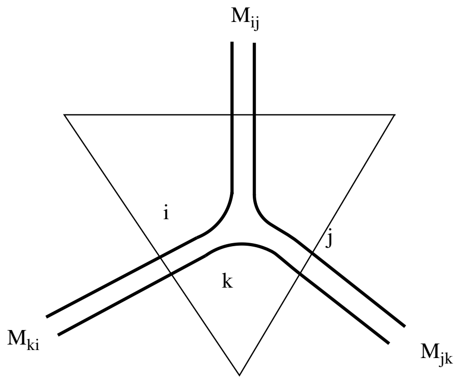
$\left(K^{-1}\right)_{jkli}$



V_{jmknli}

The composition of such building blocks is given by the tracing of indices i, j, k in the kinetic and vertex term, and represents identification of the points labeled by the same indices, in the corresponding graphical representation, and thus a sort of "higher-dimensional locality." Diagrams are then made of: (double) lines of propagation (made of two strands), non-local "vertices" of interaction (providing a rerouting of strands), faces (closed loops of strands) obtained after index contractions. The same combinatorics of indices (and thus of matrices) can be given a simplicial representation as well.





Therefore the Feynman diagrams used in evaluating the partition function $Z = \int \mathcal{D}M_{ij}\, e^{-S(M)}$ correspond to 2-dimensional simplicial complexes of arbitrary topology, because they are obtained by arbitrary gluing of edges to form triangles (in the interaction vertex) and of these triangles to one another along common edges (as dictated by the propagator). Thus a discrete 2D spacetime emerges as a virtual construction, encoding the possible interaction process of fundamentally discrete quanta of space (Figure 12.3).

We can easily compute the Feynman amplitudes for the model:

$$Z = \sum_{\Gamma} \left(\frac{g}{\sqrt{N}}\right)^{\frac{1}{2}} Z_{\Gamma} = \sum_{\Gamma} g^{V_{\Gamma}} N^{F_{\Gamma} - \frac{1}{2} V_{\Gamma}}$$

where V_{Γ} is the number of vertices of the Feynman diagram, and F_{Γ} the number of faces of the same, and N, again, the dimension of the matrices.

We ask ourselves now what is the relation of this expansion and of the Feynman amplitudes of the theory with gravity. As long as we remain at the discrete level, we can of course only expect a relation with simplicial gravity. The general idea is that each Feynman amplitude will be associated with simplicial path integrals for

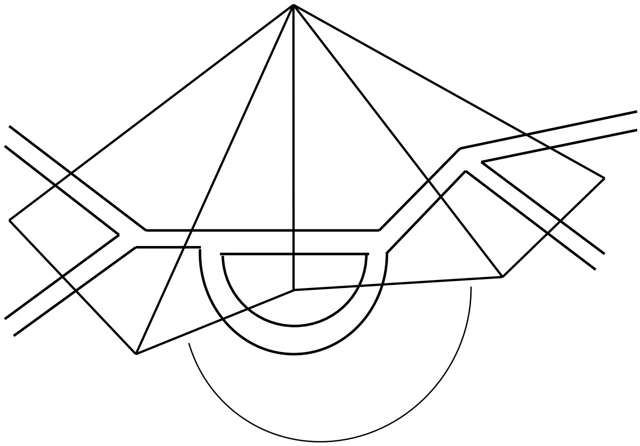


Figure 12.3 A (piece of) Feynman diagram for a matrix model, of which we give both direct and dual (simplicial) representation: the two parallel lines of propagation correspond to the two indices of the matrix; the extra line on the bottom indicates identification of the two edges of the triangles.

gravity discretized on the associated simplicial complex Δ:

$$Z_\Gamma \simeq \int \mathcal{D}g_\Delta \, e^{-S_\Delta(g)}.$$

The key to the identification of the amplitudes with a simplicial gravity path integral lies in the identity: $F_\Gamma - \frac{1}{2}V_\Gamma = v - \frac{1}{2}t = v - e + t = \chi = 2 - 2h$, where v, e, t are the numbers of vertices, edges and triangles of the simplicial complex dual to the Feynman diagram, and χ is the Euler characteristics of the same and h its genus. Thus,

$$Z = \sum_\Gamma g^{V_\Gamma} N^{\chi_\Gamma}.$$

Now, consider continuum (Riemannian) 2D GR with cosmological constant Λ, on a 2D manifold S. Its action is $\int_S d^2x \, \sqrt{g}\, (-R(g) + \Lambda) = -4\pi\, \chi \, + \, \Lambda \, A_S$, where A_S is the area of the surface. Consider then a simple discretization of the same. Chop the surface S into equilateral triangles of area a. The action will then be given by $\frac{1}{G}\int_S d^2x \, \sqrt{g}\, (-R(g) + \Lambda) \, \rightarrow \, -\frac{4\pi}{G}\, \chi \, + \, \frac{\Lambda a}{G}\, t$. Using this discretization, and defining $g = e^{-\frac{\Lambda a}{G}}$ and $N = e^{+\frac{4\pi}{G}}$, from our matrix model we get:

$$Z = \sum_\Gamma g^{V_\Gamma} N^\chi = \sum_\Delta e^{+\frac{4\pi}{G}\chi(\Delta) - \frac{a\Lambda}{G} t_\Delta}.$$

In other words, we obtain a (trivial) sum over histories of discrete GR on a given 2D complex, whose triviality is due to the fact that the only geometric variable associated with each surface is its area, the rest being only a function of topology. In addition to this sum over geometries, from our matrix model we obtain a sum over all possible 2D complexes of all topologies. In other words, the matrix model defines a discrete third quantization of GR in 2D!

The real question is now whether we can control the sum over triangulations and over topologies. The answer is in the affirmative [28, 53]: as the sum is governed only by topological parameters. Let us redefine: $M \to \frac{M}{\sqrt{N}}$ and $S(M) \Rightarrow S(M) = N\frac{1}{2}tr M^2 - N\, g\, tr M^3$. Now the Feynman expansion gives a factor N for each face (vertex of the simplicial complex), N^{-1} for each line (edge), Ng for each vertex (triangle), and thus the partition function can be recast in the form:

$$Z = \sum_{\Gamma} g^{V_\Gamma} N^{V_\Gamma - L_\Gamma + F_\Gamma} = \sum_{\Delta} g^{t_\Delta} N^{\chi(\Delta)} = \sum_{\Delta} g^{t_\Delta} N^{2-2h}$$
$$= \sum_{h} N^{2-2h} Z_h(g) = N^2 Z_0(g) + Z_1(g) + N^{-2} Z_2(g) + \cdots$$

It is then apparent that, in the limit $N \to \infty$, only spherical, trivial topology (i.e. planar, of genus 0) contributes significantly to the sum.

The second question is whether, in this limit of trivial topology, one can also define a continuum limit and match the results of the continuum 2D gravity path integral. In order to study the continuum limit, we expand $Z_0(g)$ in powers of g to obtain:

$$Z_0(g) = \sum_{V} V^{\gamma-3} \left(\frac{g}{g_c}\right)^V \simeq_{V\to\infty} (g - g_c)^{2-\gamma}$$

so that, in the limit of large number of triangles, and for the coupling constant approaching the critical value $g \to g_c$ ($\gamma > 2$), the partition function can be shown to diverge. This is a signal of a phase transition. In order to identify this phase transition as a continuum limit, we compute the expectation value for the area of a surface: $\langle A\rangle = a\,\langle t_\Delta\rangle = \langle V_\Gamma\rangle = a\frac{\partial}{\partial g}\ln Z_0(g) \simeq \frac{a}{g-g_c}$, for large V. We see that it also diverges when $g \to g_c$.

Thus we can send the area of each triangle to zero by sending the edge lengths a to zero, $a \to 0$, and simultaneously the number of triangles to infinity: $t = V \to \infty$ (continuum limit), while tuning at the same time the coupling constant to its critical value $g \to g_c$, to get finite contnuum macroscopic areas.

This defines a continuum limit of the matrix model. One can then show [28, 53] that the results obtained in this limit match those obtained with a continuum 2D gravity path integral (when this can be computed explicitly).

Let us now ask whether the third quantized framework we have (in this 2D case) can also allow us to understand and compute the contribution from non-trivial topologies in a continuum limit, and thus go beyond what can be computed in the continuum gravity path integral. The key technique is the so-called double-scaling limit [28, 53]. One can first of all show that the contribution of each given topology of genus h to the partition function is

$$Z_h(g) \simeq \sum_V V^{\frac{(\beta-2)\chi}{2}-1}\left(\frac{g}{g_c}\right)^V \simeq f_h\,(g-g_c)^{\frac{(2-\beta)\chi}{2}},$$

where the last approximation holds in the limit of many triangles (necessary for any continuum limit, and corresponding to the thermodynamics limit), and where β is a constant that can be computed.

Define now $\kappa^{-1} = N\,(g-g_c)^{\frac{(2-\beta)}{2}}$, so that we get:

$$Z \simeq \sum_h \kappa^{2h-2} f_h = \kappa^{-2} f_0 + f_1 + \kappa^2 f_2 + \cdots$$

We can then take the combined limits $N \to \infty$ and $g \to g_c$, while holding κ fixed. The result of this double limit is a continuum theory to which all topologies contribute.

The area of matrix models is vast and rich of results, not only in the 2D quantum gravity context, but ranging from condensed matter physics to hot topics in mathematical physics, from string theory to mathematical biology. For all of this, we can only refer to the literature.

Tensor models?

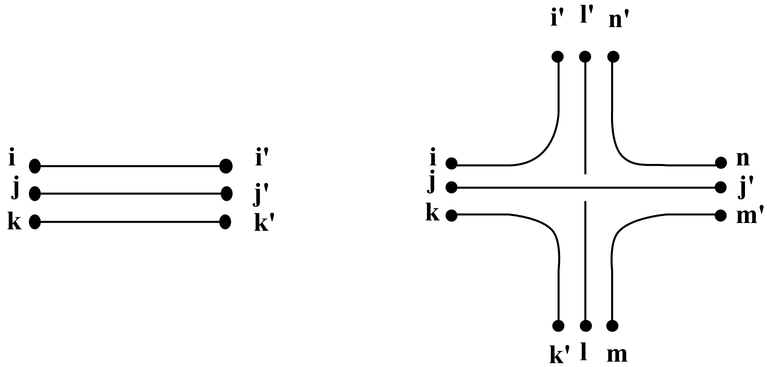
Now we generalize further in combinatorial dimension: from 2D to 3D. This is achieved by going from 1D objects, represented graphically as edges, and mathematically by matrices, to 2D objects, represented graphically as triangles and mathematically by tensors [2, 56]. Obviously, any 2D structure (squares, polygons, etc.) would be a possible choice, but any other 2D structure could be built out of triangles, so we stick to what looks like the most fundamental choice. This also means that, in the Feynman expansion of the corresponding theory, we expect to generalize from 2D simplicial complexes to 3D ones. We then define $N \times N \times N$ tensors T_{ijk}, with $i, j, k = 1, \ldots, N$ and an action for them given by

$$S(T) = \frac{1}{2} trT^2 - \lambda\, trT^4 = \frac{1}{2}\sum_{i,j,k} T_{ijk}T_{kji} - \lambda \sum_{ijklmn} T_{ijk}T_{klm}T_{mjn}T_{nli},$$

where the choice of combinatorics of tensors in the action, and of indices to be traced out, is made so as to represent, in the interaction term, four triangles (tensors) glued pairwise along common edges (common indices) to form a closed tetrahedron (3-simplex). The kinetic term dictates the gluing of two such tetrahedra along common triangles, by identification of the edges. From the action above we read out the kinetic and vertex terms:

$$K_{ijki'j'k'} = \delta_{ii'}\delta_{jj'}\delta_{kk'} = (K^{-1})_{ijki'j'k'},$$
$$V_{ii'jj'kk'll'mm'nn'} = \delta_{ii'}\delta_{jj'}\delta_{kk'}\delta_{ll'}\delta_{mm'}\delta_{nn'},$$

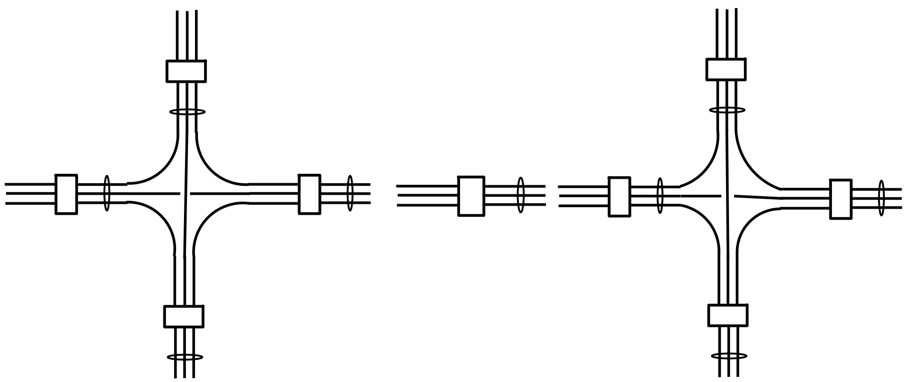
which can be represented graphically as



and then used to expand the partition function perturbatively:

$$Z = \int \mathcal{D}T \, e^{-S(T)} = \sum_{\Gamma} \lambda^{V_\Gamma} Z_\Gamma.$$

Feynman diagrams are again obtained by contraction of vertices with propagators over internal indices:



By construction, Feynman diagrams are again formed by vertices, lines and faces, but now also form "bubbles" (3-cells), and are dual to 3D simplicial complexes. See Figure 12.4.

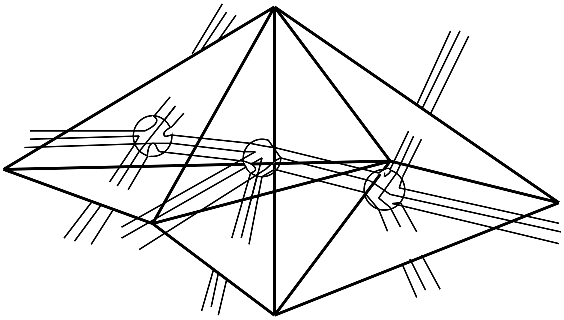


Figure 12.4 A (piece of) Feynman diagram for a tensor model, of which we give both direct and dual (simplicial) representation; the three parallel lines of propagation (dual to the three edges in the triangles of the simplicial complex) correspond to the three indices of the tensor.

As a result, Z is defined as a sum over all 3D simplicial complexes including manifolds as well as pseudo-manifolds (i.e. singular complexes such that the neighborhood of some points is not homeomorphic to a 3-ball), because we impose *a priori* no restriction on the gluing procedure of vertices by means of propagators.

Do these models provide a good definition of 3D quantum gravity? The answer, unfortunately, is a resounding no. None of the nice features of matrix models export to the case of tensor models. First of all there is no direct/nice relation between the Feynman amplitudes of the above tensor model with 3D simplicial (classical or quantum) gravity. Even though 3D gravity is a topological theory, with no local propagating degrees of freedom and only locally flat solutions (in the absence of a cosmological constant), it is still a highly non-trivial theory. The amplitudes of tensor models are too simple to capture either the flatness of geometry or the topological character of the quantum gravity partition function (as for example does Chern–Simons theory). They do not have enough data in the amplitudes associated with each simplicial complex, or in boundary states. Second, there is no way to separate the contribution of manifolds from that of pseudo-manifolds, i.e. to suppress singular configurations or even to identify them clearly. Third, the expansion in sum over simplicial complexes cannot be organized in terms of topological invariants, and so there is no control over the topology of the diagrams summed over. Also the last two issues can be thought to be due to the lack of data and structure in the Feynman amplitudes of the theory.

Clearly, the process of combinatorial generalization can be continued to higher tensor models whose Feynman diagrams will be higher simplicial complexes. It is clear, however, that the difficulties encountered with 3D tensor models are not going to be solved magically if we do not render the structure of the corresponding quantum amplitudes richer and more interesting. In particular, just as we do in the 1D case, i.e. in the case of particles, we could generalize matrix and tensor models in the direction of adding degrees of freedom, i.e. defining corresponding field theories. In the process, the indices of the tensor models will be replaced by variables living in appropriate domain spaces, and sums over indices by appropriate sums or integrals over these domain spaces, while maintaining their combinatorial pairing in the action. This pairing will make the resulting field theories *combinatorially non-local*, as we anticipated group field theories are. In fact this is in many ways the defining property of group field theories.

The prototype for a field theory of this non-local type and for a choice of domain space D would be, for the 2D case:

$$S(\phi) = \frac{1}{2}\int_D [dg]\phi(g_1, g_2)\phi(g_2, g_1) - \frac{\lambda}{3!}\int [dg]\,\phi(g_1, g_2)\phi(g_2, g_3)\phi(g_3, g_1),$$

with appropriate integrations over the domain space D, and the same identification of field arguments as in the indices of the matrix model, while for the field-theoretic generalization of the tensor model for T_{ijk} we get:

$$\begin{aligned} S(\phi) = &\frac{1}{2}\int_D [dg]\phi(g_1, g_2, g_3)\phi(g_3, g_2, g_1) \\ &- \frac{\lambda}{4!}\int [dg]\,\phi(g_1, g_2, g_3)\phi(g_3, g_4, g_5)\phi(g_5, g_6, g_1)\phi(g_6, g_4, g_2). \end{aligned} \tag{12.3}$$

The definition of good group field theory models for quantum gravity, of course, would require a careful choice of domain space and of classical action (kinetic and vertex functions). Here is where the input from other approaches is crucial, and where their characteristic structures are incorporated into the GFT formalism. In particular, the domain space of GFTs for quantum gravity in three and higher dimensions is given by either a group manifold (from which the name of the formalism and the choice of notation in the examples of actions is given above) or, more recently, by the corresponding Lie algebras. This allows also a rewriting of the GFT action (and amplitudes) in terms of group representations, as we will see. We now motivate this choice(s) for the domain space.

12.1.7 Ingredients from loop quantum gravity and simplicial topological theories

Let us first look at loop quantum gravity. As we mentioned already, after passing to connection variables valued in the Lorentz algebra $\mathfrak{so}(3, 1)$ and suitable frame-fixing of the tetrad variables to $\mathfrak{su}(2)$ tetrads, GR becomes similar to a gauge theory for the gauge group SU(2), and as such is quantized in loop quantum gravity, taking of course particular care of the diffeomorphism invariance that characterizes the theory [90]. Due to the gauge-fixing, therefore, what would be initially an SO(3, 1) gauge theory is reduced to an SU(2) one, with a similar reduction of the conjugate variables to the connection. Let us see briefly (and simplifying considerably the construction) how the kinematical phase space of the theory is defined.

If we assume the SU(2)-bundle to be trivial, then every SU(2)-connection can be seen as an $\mathfrak{su}(2)$-valued one-form on the 3-dimensional base manifold Σ. Being a one-form A can naturally (i.e. without referring to a background metric) be integrated along 1-dimensional submanifolds of Σ, namely along embedded edges e:

$$\int_e A := \int_e A_a^j \tau_j dx^a. \tag{12.4}$$

The conjugate variable to the connection, the triad E, being an $\mathfrak{su}(2)$-valued vector density, has a natural associated 2-form $(*E)^j_{ab}(\sigma) = \varepsilon_{abc} E^c_j(\sigma)$. This 2-form can be integrated along submanifolds of codimension one, namely analytic 2-surfaces S (by means of appropriate parallel transports): $E(S, f) := \int_S (*E)^j f_j$. Here f is a smearing function with values in $\mathfrak{su}(2)^*$, the topological dual of $\mathfrak{su}(2)$ and f_j are its components in a local basis.

To get quantities with a nicer behavior under SU(2)-transformations one introduces the holonomy

$$h_e(A) := \mathcal{P} \exp\left[-\int_e A\right], \tag{12.5}$$

where $\mathcal{P}$ denotes the path-ordering symbol, which is of course an SU(2) group element. For a graph γ with $|\gamma|$ edges the holonomy assigns an element $h_e(A) \in$ SU(2) to every edge. One defines the space

$$\mathrm{Cyl}^\gamma = \{C^\gamma : \mathcal{A} \to \mathbb{C}; A \mapsto C^\gamma(A) \,|\, C^\gamma(A) := c(h_{e_1}(A), h_{e_2}(A), \dots h_{e_{|\gamma|}}(A))\} \tag{12.6}$$

of functions called *cylindrical with respect to* γ that depend on A only through the holonomies $h_{e_i}(A)$ and $c : \mathrm{SU}(2)^{|\gamma|} \to \mathbb{C}$ is a continuous complex valued function.

The configuration space of the theory is defined to be the space Cyl (without reference to a specific graph γ) as the space of functions that are cylindrical with respect to *some* graph. One can then show [90] that the fluxes $E(S, f)$ are vector fields on Cyl.

The classical Poisson algebra between cylindrical functions (including single holonomies) can be computed in full generality, i.e. for arbitrary graphs γ and surfaces S. The basic feature is that holonomies Poisson commute, fluxes and holonomies have non-zero commutators depending on the intersection points between graphs γ to which holonomies are associated and surfaces S on which the fluxes are smeared, while fluxes associated with any two surfaces (including coincident ones) do not commute. The non-commutativity of the fluxes even at the classical level is crucial for what follows. The algebra is in general rather complicated, to the point that the general commutator between fluxes is not known. It simplifies considerably if one considers an alternative definition of the fluxes defined only on "elementary" surfaces with single intersection points with the graph (this way, one has to label each state by both graph and a set of dual surfaces). The fluxes are defined as

$$E_i^e = tr[\tau_i \int_{S_e} \mathrm{Ad}(h_{e,x}) E(x)], \tag{12.7}$$

where $h_{e,x}$ is the holonomy along the path from the starting point of the edge e to the point x on the surface x.

In fact, considering a single link e and a single elementary dual surface, also labeled e, the phase space of the theory reduces to the cotangent bundle of SU(2), $T^*\mathrm{SU}(2)$, with fundamental variables being the holonomy h_e along the link e and the dual triad E_e^i and fundamental Poisson brackets being

$$\begin{aligned}
\{h_e, h'_e\}_\gamma &= 0, \\
\{E_i^e, h'_e\} &= \delta^e_{e'} \frac{\tau_i}{2} h_e, \\
\{E_i^e, E_j^{e'}\} &= -\delta^{ee'} \varepsilon_{ijk} E_k^e,
\end{aligned} \tag{12.8}$$

with τ_i the generators of the $\mathfrak{su}(2)$ Lie algebra. The last bracket among fluxes is clearly the $\mathfrak{su}(2)$ bracket.

Now, interestingly, this is also the phase space of discrete topological BF theory [32], with SU(2) as gauge group, and thus sets the basic kinematical stage for the simplicial path integral quantization of that theory. This fact is going to be crucial in the following, when discussing the GFT model for 3D gravity (which in 3D coincides with BF theory). Also simplicial BF theory, then, is based on a configuration space given by cylindrical functions for the gauge group, the difference with LQG (or the covariant version of the same) is therefore only in the constraints that one imposes on such functions to implement the dynamics.

12.1.8 Some mathematical tools

Given this phase space, and considering for now a single edge of any graph labeling the states, there is a natural Fourier transform that can be introduced and used to map between configuration and "momentum" space. This is the so-called non-commutative "group Fourier transform," introduced first in [43], and whose properties have been analyzed in [49, 62, 66]. Introduced first in the context of spin foam models, it has recently been used extensively in the GFT context [9, 12], and then applied also in LQG [13] to give a presentation of the theory in terms of metric (flux) variables. We introduce it briefly here, and then we will show its application in the GFT model for 3D gravity.

The group Fourier transform is based on the plane waves

$$e : \mathrm{SU}(2) \times \mathfrak{su}(2) \to \mathbb{C}; (g, x) \mapsto e_g(x) := e^{itr(xg)} \tag{12.9}$$

where $x := \vec{x} \cdot \vec{\tau}$ in a basis τ^i of $\mathfrak{su}(2)$ and the trace is taken in the fundamental representation. One can always identify $\mathfrak{su}(2)$ with $\mathbb{R}^3$ as vector spaces and thus e_g can be interpreted as elements of $C(\mathbb{R}^3)$. Denote the closure of the linear span of these elements as $C_\kappa(\mathbb{R}^3)$. We can then introduce a non-commutative product on the algebra of functions $C_\kappa(\mathbb{R}^3)$, starting from plane waves

$$* : C_\kappa(\mathbb{R}^3) \times C_\kappa(\mathbb{R}^3) \to C_\kappa(\mathbb{R}^3); (e_{g_1}, e_{g_2}) \mapsto e_{g_1} * e_{g_2} := e_{g_1 g_2}, \tag{12.10}$$

and extending it to all of $C_\kappa(\mathbb{R}^3)$ by linearity. $C_\kappa(\mathbb{R}^3)$ endowed with this non-commutative product is referred to as $C_{*,\kappa}(\mathbb{R}^3)$.

Using these plane waves and the Haar measure dg on SU(2) one defines the group Fourier transform as

$$F : C(\mathrm{SU}(2)) \to C_{*,\kappa}(\mathbb{R}^3); f(g) \mapsto \widetilde{f}(x) := \int dg e_g(x) f(g). \tag{12.11}$$

As defined above, the Fourier transform is not invertible, but can be made so by suitable modification of the plane waves, as shown in [49, 62, 66]. We do not indicate explicitly this modification (which basically amounts to multiplication by a polarization vector, keeping track of the hemisphere of SU(2) on which the group element g belongs) in order not to clutter the notation. Therefore, using the star product (12.10), and these (modified) plane waves, we get a bijection

$$F : C(\mathrm{SU}(2)) \to C_{*,\kappa}(\mathbb{R}^3); f(g) \mapsto f(x) := \int dg e_g(x) f(g), \tag{12.12}$$

$$F^{-1} : C_{*,\kappa}(\mathbb{R}^3) \to C(\mathrm{SU}(2)); f(x) \mapsto f(g) := \int dx (e_g * f)(x), \tag{12.13}$$

which maps functions on SU(2) onto ordinary functions living on $\mathbb{R}^3$ (equivalently, on the $\mathfrak{su}(2)$ Lie algebra). The non-commutativity of SU(2) (and of its Lie algebra) is taken into account via the star product (12.10). The extension of this group Fourier transform to functions of arbitrary numbers of groups or Lie algebra elements, and to arbitrary groups, will be a crucial element of the GFT formalism, as it provides a duality of representations for the GFT field as a function of group elements or of Lie algebra elements. Clearly, it also plays an important role in dealing with simplicial BF theory. In fact, the Feynman amplitudes of the GFT we will present, in the Lie algebra basis, will be given by the simplicial path integral for 3D SU(2) BF theory (or 3D gravity).

For arbitrary (square integrable) functions on groups, another type of generalization of the usual Fourier transform on $\mathbb{R}^d$ is available. For compact groups (like the rotation group in any dimension), this is given by the Peter–Weyl decomposition of the function itself into irreducible representations. For SU(2) it gives:

$$f(g) = \sum_j (2j+1) f^j_{mn} D^j_{mn}(g),$$

where (with repeated indices summed over) $j \in \mathbb{N}/2$ labels the irreducible representations of SU(2), m, n are indices labeling a basis in the vector space of the representation j, and D^j_{mn} are the Wigner representation matrices, here playing the role of plane waves. This maps functions of group elements to functions of representation labels, and can be inverted to give:

$$f^j_{mn} = \int dg\, f(g) \left(D^j_{mn}(g) \right)^* .$$

The extension of this decomposition to cylindrical functions is the basis of the spin network representation of loop quantum gravity, in which basis states correspond to graphs γ whose links are labeled by irreducible representations of SU(2), and vertices by intertwiners of the group, after imposition of gauge invariance at the same vertices (Figure 12.5). Similarly, the spin network basis provides an alternative representation for the GFT field and for the corresponding states. In this representation, as we will see, the field is represented not as a fundamental simplex but as a single spin network vertex, the building block for the construction of arbitrary spin networks.

12.1.9 The spin foam idea

A covariant path integral quantization of a theory based on spin networks will have as histories a higher-dimensional analogue of them: a spin foam [72, 74, 81], i.e. a 2-complex (collection of faces bounded by links joining at vertices) with representations of the Lorentz (or SU(2)) group attached to its faces, in such a way that any slice or any boundary of it, corresponding to a spatial hypersurface, will be given by a spin network (Figure 12.6). Spin foam models [72, 74, 81] are intended

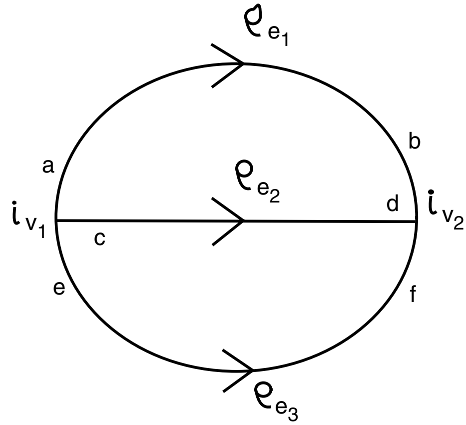


Figure 12.5 A spin network.

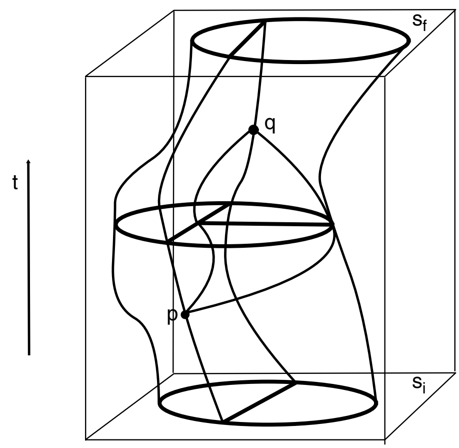


Figure 12.6 A spin foam.

to give a path integral quantization of gravity based on these purely algebraic and combinatorial structures.

In most of the current models, the combinatorial structure of the spin foam is restricted to be topologically dual to a simplicial complex of appropriate dimension, so that to each spin foam 2-complex corresponds a simplicial spacetime, with the representations attached to the 2-complex providing quantum geometric information to the simplicial complex. The models are then defined by an assignment of a quantum probability amplitude (here factorized in terms of face, edge, and vertex contributions) to each spin foam σ summed over, depending on the representations

ρ labeling it, also being summed over:

$$Z = \sum_{\sigma} w(\sigma) \sum_{\{\rho\}} \prod_{f} A_f(\rho_f) \prod_{e} A_e(\rho_{f|e}) \prod_{v} A_v(\rho_{f|v}).$$

One then has an implementation of a sum-over-histories for gravity in a combinatorial–algebraic context. We will show that this spin foam representation is characteristic of the GFT Feynman amplitudes, and that it is dual to the representation of the same in the form of a simplicial path integral, a duality stemming from the above duality of representations for functions on group manifolds. A multitude of results have already been obtained in the spin foam approach, for which we refer to [72, 74, 81].

Let us summarize the ingredients and aims of the GFT formalism, before entering the details of one specific GFT model. We want to define a quantum field theory of fundamental building blocks of quantum space, whose combinations can build up arbitrary spatial topological manifolds, and whose dynamics and interaction processes generate arbitrary spacetime topologies, thanks to a peculiar non-locality of field pairing in the interaction term of the GFT classical action. They are thus a sort of discrete or finitary, and local, third quantization of gravity. In this, they represent a generalization of matrix models to arbitrary dimension. The arguments of the GFT field are given either as group elements or as Lie algebra elements or as group representations. In this, GFTs incorporate the kinematical description of the geometry of loop quantum gravity and discrete topological BF theories. The Feynman amplitudes of the theory are, as we will see, given by simplicial path integrals or, equivalently, by spin foam models.

12.2 Dynamics of 2D quantum space as a group field theory

We present here in some detail the construction and perturbative analysis of the GFT model for 3D Riemannian gravity, first introduced, in the group picture, by Boulatov [24], whose expansion in group representations gives the Ponzano–Regge spin foam model [47], recently reformulated in terms of non-commutative Lie algebra variables in [12].

12.2.1 The kinematics of quantum 2D space in GFT: quantum simplices and spin networks

Consider a triangle in $\mathbb{R}^3$. We consider its (second quantized) kinematics to be encoded in the GFT field φ. We work here with real fields for simplicity only. The GFT field can be understood as living on the space of possible geometries for the triangle itself, or on the corresponding conjugate space. We parametrize the possible geometries for the triangle in terms of three $\mathfrak{su}(2)$ Lie algebra elements attached to its three edges. These are to be thought of as fundamental triad variables obtained by discretization of continuum triad fields along the edges of the same triangle in line with the LQG and discrete BF construction.

The field is then a function

$$\varphi : (x_1, x_2, x_3) \in \mathfrak{su}(2)^3 \longrightarrow \varphi(x_1, x_2, x_3) \in \mathbb{R}.$$

We do not assume any symmetry of the field under permutation of the arguments. Different choices are possible, as the field can be taken to be in any representation of the permutation group acting on its arguments. The choice of representation made will influence the type of combinatorial complexes generated as Feynman diagrams of the theory [79]. This will not concern us here. We will also see that a simple modification of the construction, defining "colored GFTs", can be used to make this choice somewhat irrelevant from the point of view of the same combinatorics.

Using the non-commutative group Fourier transform [43, 49, 62] introduced earlier, the same GFT field can be recast as a function of SU(2) group elements.

Recapitulating and detailing a bit more of its definition, this transform stems from the definition of plane waves $\mathrm{e}_g(x) = e^{i\vec{p}_g\cdot\vec{x}}$ as functions on $\mathfrak{g} \sim \mathbb{R}^n$, depending on a choice of coordinates $\vec{p}_g = \mathrm{Tr}\, g\vec{\tau}$ on the group manifold, where are i times the Pauli matrices and Tr is the trace in the fundamental representation.[1] For $x = \vec{x}\cdot\vec{\tau}$ and $g = e^{\theta\vec{n}\cdot\vec{\tau}}$, we thus have

$$\mathrm{e}_g(x) = e^{i\mathrm{Tr}\,xg} = e^{-2i\sin\theta\vec{n}\cdot\vec{x}}.$$

The image of the Fourier transform of functions on SU(2) inherits an algebra structure from the convolution product on the group, given by the $\star$-product defined on plane waves as

$$\mathrm{e}_{g_1} \star \mathrm{e}_{g_2} = \mathrm{e}_{g_1 g_2}.$$

Fourier transform and $\star$-product extend straightforwardly to functions of several variables like the GFT field (and generic cylindrical functions), so that

$$\varphi(x_1, x_2, x_3) = \int [dg]^3\, \varphi(g_1, g_2, g_3)\, \mathrm{e}_{g_1}(x_1)\mathrm{e}_{g_2}(x_2)\mathrm{e}_{g_3}(x_3)$$

and the GFT field can also be seen as a function of three group elements, thought of as parallel transports of the gravity connection along fundamental links dual to the edges of the triangle represented by φ, and intersecting them only at a single point. Notice once more the consistency with the LQG construction.

[1] We are neglecting once more, for notational simplicity, the subtlety in the definition of the plane waves, concerning the need to keep track of the position on the upper or lower SU(2) hemisphere of the group element g; for details see [10, 12, 49, 62, 66].

In order to define a geometric triangle, the vectors (Lie algebra elements) associated with its edges cannot be independent. Indeed, they have to "close" to form a triangle, i.e. they have to sum to zero. We thus impose a constraint on the field:

$$\varphi = C \star \varphi, \quad C(x_1, x_2, x_3) = \delta_0(x_1 + x_2 + x_3)$$

by means of the projector C, where δ_0 is the element $x = 0$ of the family of functions

$$\delta_x(y) := \int dg\, \mathrm{e}_{g^{-1}}(x) \mathrm{e}_g(y).$$

These play the role of Dirac distributions in the non-commutative setting, in the sense that[2]

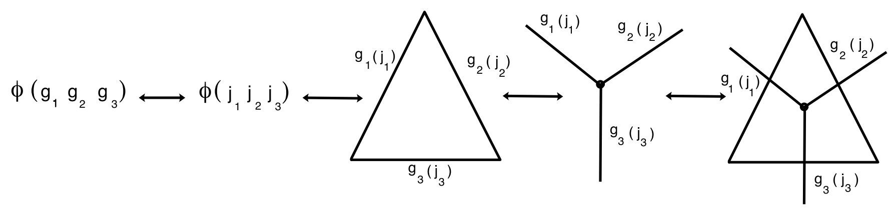
$$\int d^3y\, (\delta_x \star f)(y) = \int d^3y\, (f \star \delta_x)(y) = f(x).$$

One can also show that $(\delta_x \star f)(y) = (\delta_y \star f)(x)$.

In terms of the dual field $\varphi(g_1, g_2, g_3)$, the same closure constraint implies invariance under the diagonal (left) action of the group SU(2) on the three group arguments, imposed by projection P:

$$\varphi(g_1, g_2, g_3) = P\varphi(g_1, g_2, g_3) = \int_{\mathrm{SU}(2)} dh\, \phi(hg_1, hg_2, hg_3). \tag{12.14}$$

Because of this gauge invariance, which is in fact imposed in the same way as the Gauss constraint is imposed on cylindrical functions in LQG, the field can best be depicted graphically as a 3-valent vertex with three links, dual to the three edges of the closed triangle:



This object, will be the GFT building block of quantum space.

One obtains another representation of the GFT field by means of Peter–Weyl decomposition into irreducible representations, in the same way as one obtains the spin network expansion of generic cylindrical functions in LQG.

[2] Although behaving like a proper delta distribution with respect to the $\star$ product, under integration, this is a regular function when seen as a function on $\mathbb{R}^3$. In fact, seen as a function of $\mathbb{R}^3$, δ_0 is the regular function peaked on $x = 0$ given by $\delta_0(x) \propto J_1(|x|)/|x|$, where J_1 is the first Bessel function [49, 62, 66].

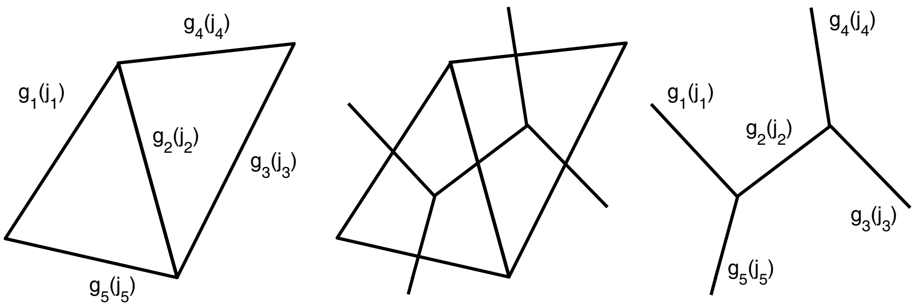
The invariant field decomposes in SU(2) representations as:

$$\varphi(g_1, g_2, g_3) = \sum_{j_1, j_2, j_3} \varphi^{j_1 j_2 j_3}_{m_1 m_2 m_3} D^{j_1}_{m_1 n_1}(g_1) D^{j_2}_{m_2 n_2}(g_2) D^{j_3}_{m_3 n_3}(g_3) C^{j_1 j_2 j_3}_{n_1 n_2 n_3}, \quad (12.15)$$

where $C^{j_1 j_2 j_3}_{n_1 n_2 n_3}$ is the Wigner-invariant 3-tensor, the 3j-symbol.

This expansion in irreducible group representations can be understood as a representation of the GFT field in terms of the quantum numbers associated with the quantized geometry of the triangle it represents. The j_i's label eigenvalues of the length operators corresponding to its edges, while the angular momentum indices encode directional degrees of freedom. This is confirmed by canonical analysis of the quantum geometry of the triangle, as well as by geometric quantization methods [6, 14, 44].

Multiple fields can be convoluted (in the group or Lie algebra picture) or traced (in the representation picture) with respect to some common argument. This represents the gluing of multiple triangles along common edges, and thus the formation of more complex simplicial structures, or, dually, of more complicated graphs:



The corresponding field configurations thus represent extended chunks of quantum space, or many-GFT-particle states. Generic polynomial GFT observables would be given by this type of construction, and thus be associated with a particular quantum space. This includes, of course, any open configuration, in which the arguments of the involved GFT fields are not all convoluted or contracted, representing a quantum space (not necessarily connected) with boundary.

We would like to point out, that, as in the case of tensor models, combinatorial generalizations can be considered, since there is no *a priori* restriction on how many arguments a GFT field can have. Once closure constraint or gauge invariance has been imposed on such a generalized field with n arguments, it can be taken to represent a general n-polygon (dual to an n-valent vertex), and glued to another in order give a polygonized quantum space in the same way as we have outlined for triangles.

12.2.2 Classical (third quantized) dynamics of 2D space in GFT

We now define a classical dynamics for the introduced GFT field. The prescription for the interaction term, as in tensor models, is simple: four geometric triangles should be glued to one another, along common edges, to form a 3-dimensional geometric tetrahedron. The kinetic term should encode the gluing of two tetrahedra along common triangles, by identification of their edge variables. There is no other dynamical requirement at this stage.

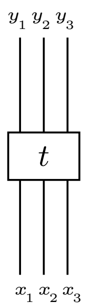
Thus we take four fields and identify pairwise their edge Lie algebra elements (triad edge vectors), with the combinatorial pattern of the edges of a tetrahedron; this gives the potential term in the action, weighted by an arbitrary coupling constant. And we take two more fields and identify their arguments; this defines the kinetic term in the action. Naming $\varphi_{123} = \varphi(x_1, x_2, x_3)$, the combinatorial structure of the action is then

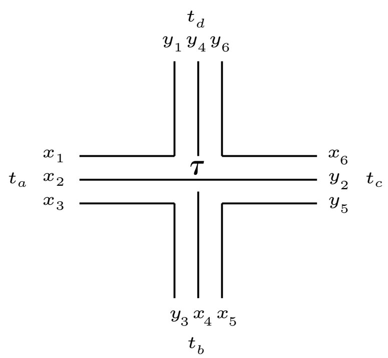
$$S = \frac{1}{2}\int [dx]^3\, \varphi_{123} \star \varphi_{123} - \frac{\lambda}{4!}\int [dx]^6\, \varphi_{123} \star \varphi_{345} \star \varphi_{526} \star \varphi_{641},$$

where it is understood that $\star$-products relate repeated indices as $\phi_i \star \phi_i := (\phi \star \phi_-)(x_i)$, with $\phi_-(x) = \phi(-x)$.

Notice once more that there is no difficulty, as there is none in matrix models, in defining or dealing with a combinatorial generalization of the action in which, for given building block $\varphi(x_i)$, one adds other interaction terms corresponding to the gluing of triangles (or polygons) to form general polyhedra or even more pathological configurations (e.g. with multiple identifications among triangles). The only restriction may come from the symmetries of the action.

The structure of this action is best visualized in terms of diagrams, similar to those used in the discussion of tensor models. Kinetic and interaction terms identify a propagator and a vertex given by:





$$\int dh_t \prod_{i=1}^{3} (\delta_{-x_i} \star \mathrm{e}_{h_t})(y_i), \qquad \int \prod_t dh_t \prod_{i=1}^{6} (\delta_{-x_i} \star \mathrm{e}_{h_{tt'}})(y_i), \tag{12.16}$$

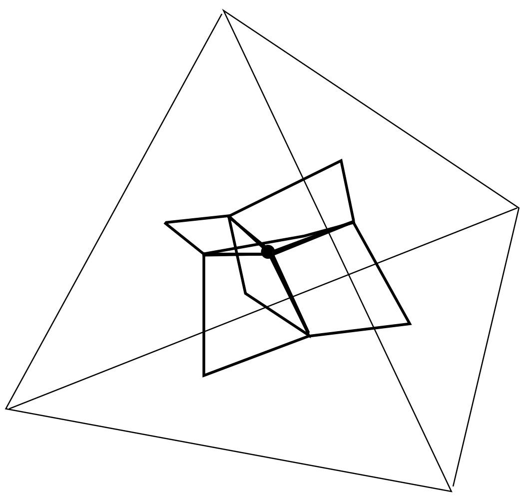
with $h_{tt'} := h_{t\tau}h_{\tau t'}$, where we have used "$t$" for triangle and "$\tau$" for tetrahedron. The group variables h_t and $h_{t\tau}$ arise from (12.14), and should be interpreted as parallel transports through the triangle t for the former, and from the center of the tetrahedron τ to triangle t for the latter.

The integrands in (12.16) factorize into a product of functions associated with strands (one for each field argument), with a clear geometrical meaning: the pair of variables (x_i, y_i) associated with the same edge i corresponds to the edge vectors seen from the frames associated with the two triangles t, t' sharing it. The vertex functions state that the two variables are identified, up to parallel transport $h_{tt'}$, and up to a sign labeling the two opposite edge orientations inherited by the triangles t, t'. The propagator encodes a similar gluing condition, allowing for the possibility of a further mismatch between the reference frames associated with the same triangle in two different tetrahedra. In this non-commutative Lie algebra representation of the field theory, the geometric content of the action is particularly transparent.

Using the group Fourier transform we can obtain a pure group representation of the theory (the original definition of it by Boulatov):

$$S_{3d}[\phi] = \frac{1}{2}\int [dg]^3 \varphi(g_1, g_2, g_3)\varphi(g_3, g_2, g_1)$$
$$- \frac{\lambda}{4!}\int [dg]^6 \varphi(g_1, g_2, g_3)\varphi(g_3, g_4, g_5)\varphi(g_5, g_2, g_6)\varphi(g_6, g_4, g_1). \tag{12.17}$$

The combinatorics of arguments in the action has already been discussed. It is also illustrated below, for a single tetrahedron (interaction term), where one can see both the tetrahedron, on whose edges are associated Lie algebra elements, and its topological dual, on whose boundary links are associated conjugate group elements:



Labeling with indices $i, j = 1, \ldots, 4$ the triangles in each tetrahedron, and thus the group elements h imposing left diagonal invariance, and by a pair (ij) the edges shared by the triangles i and j, and thus the arguments of the field g_{ij} and g_{ji} associated with the same edge in the triangles i and j, the kinetic and vertex functions are:

$$\mathcal{K}(g_i, \tilde{g}_i) = \int dh \prod_{i=1}^{3} \delta(g_i h \tilde{g}_i^{-1}), \quad \mathcal{V}(g_{ij}) = \prod_{i=1}^{4} \int dh_i \prod_{i \neq j} \delta(g_{ij} h_i h_j^{-1} g_{ji}^{-1}). \tag{12.18}$$

Also in these sets of variables, the geometric content of the model can be read out rather easily: the six delta functions in the vertex term encode the flatness of each "wedge," i.e. of the portion of each dual face inside a single tetrahedron [74, 81]. This flatness is characteristic of the piecewise-flat context in which the GFT models are best understood.

It is clear that the result of the calculation of the Feynman amplitudes in this representation will be a pure SU(2) gauge theory for a discrete connection associated with the various elements of the simplicial complex dual to each Feynman diagram, and to the Feynman diagram itself. We will show the form of the Feynman amplitudes explicitly below.

Before doing so, however, we also present the form of the action in representation space, obtained after Peter–Weyl decomposition of the GFT field.

Using the expansion of the field (12.15), performing the appropriate contractions and using standard properties of the $3j$-symbols from SU(2) recoupling theory, the action becomes:

$$\begin{aligned} S(\varphi) = &\frac{1}{2} \sum_{\{j\},\{m\}} \varphi^{j_1 j_2 j_3}_{m_1 m_2 m_3} \varphi^{j_3 j_2 j_1}_{m_3 m_2 m_1} \\ &- \frac{\lambda}{4!} \sum_{\{j\},\{m\}} \varphi^{j_1 j_2 j_3}_{m_1 m_2 m_3} \varphi^{j_3 j_4 j_5}_{m_3 m_4 m_5} \varphi^{j_5 j_2 j_6}_{m_5 m_2 m_6} \varphi^{j_6 j_4 j_1}_{m_6 m_4 m_1} \left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{array} \right\}, \end{aligned} \tag{12.19}$$

from which one reads the kinetic and vertex terms:

$$\begin{aligned} \mathcal{K} = &= \mathcal{K}^{-1} = \delta_{j_1 \tilde{j}_1} \delta_{m_1 \tilde{m}_1} \delta_{j_2 \tilde{j}_2} \delta_{m_2 \tilde{m}_2} \delta_{j_3 \tilde{j}_3} \delta_{m_3 \tilde{m}_3} \\ \mathcal{V} &= \delta_{j_1 \tilde{j}_1} \delta_{m_1 \tilde{m}_1} \delta_{j_2 \tilde{j}_2} \delta_{m_2 \tilde{m}_2} \delta_{j_3 \tilde{j}_3} \delta_{m_3 \tilde{m}_3} \delta_{j_4 \tilde{j}_4} \delta_{m_4 \tilde{m}_4} \delta_{j_5 \tilde{j}_5} \delta_{m_5 \tilde{m}_5} \delta_{j_6 \tilde{j}_6} \delta_{m_6 \tilde{m}_6} \\ &\times \left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{array} \right\}, \end{aligned}$$

where $\Delta_j = 2j + 1$ is the dimension of the representation j and for each vertex of the 2-complex we have a so-called $6j$-symbol, a real invariant (under group action on the representation spaces) function of the six representations meeting at that vertex. The geometry behind the construction is rather obscure in this formulation. However, it gives the dynamics directly in terms of quantum numbers labeling the states of the theory, thus allowing us to deal more easily with the quantum properties of the model.

Notice also that the vertex amplitude could have been obtained simply by taking four SU(2) intertwiners (one per triangle in the boundary of a tetrahedron), associated with spin network vertices with links labeled by representations of SU(2), which are the known quantum states of BF theory, and gluing them (tracing over corresponding magnetic indices) pairwise along links, with the combinatorial pattern of the edges of the tetrahedron. This type of construction is at the root of the definition of several current spin foam models for 4D gravity.

The classical dynamics of GFT models is basically unknown territory (with the exception of the few results in [38, 54] that we will discuss briefly at the end of this chapter). The classical equations of motion for this model are, in group space:

$$\int dh\, \phi(g_1 h, g_2 h, g_3 h) - \frac{\lambda}{3!}\int [dh_i] \prod_{j=4}^{6}\int dg_j\, \phi(g_3 h_1, g_4 h_1, g_5 h_1)$$
$$\phi(g_5 h_2, g_6 h_2, g_2 h_2)\phi(g_6 h_3, g_4 h_3, g_1 h_3) = 0$$

but can of course be written both in Lie algebra space, where they also look like complicated integral equations, as well as in the form of purely algebraic equations in representation space.

The role and importance of these equations from the point of view of the GFT per se are obvious: they define the classical dynamics of the field theory, they would allow the identification of classical background configurations and non-trivial GFT phases, etc. From the point of view of quantum gravity, considering the interpretation of the GFT as a discrete "third quantization" of gravity, these classical GFT equations encode fully the quantum dynamics of the underlying (simplicial) canonical quantum gravity theory, or equivalently the quantum dynamics of first quantized spin networks, thus implementing both Hamiltonian and diffeomorphism constraints. This is in analogy with, for example, the Klein–Gordon equation, which represents at the same time the classical equation of motion of a (free) scalar field theory and the full quantum dynamics for the corresponding first quantized (free) theory.

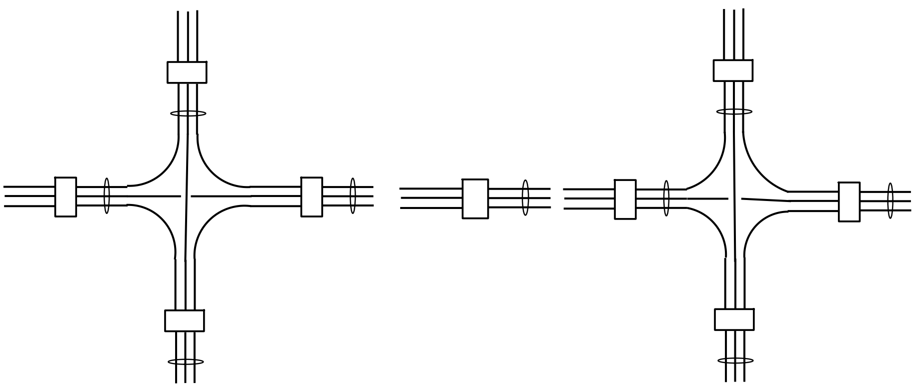
12.2.3 Quantum dynamics of 2D space in GFT: duality of simplicial path integrals and spin foam models

Let us now turn to the quantum dynamics of group field theories. In absence of a non-perturbative definition, this is defined by the perturbative expansion of the partition function in Feynman diagrams. As we have discussed, these Feynman diagrams are by construction cellular complexes topologically dual to simplicial complexes. Here we obtain 3-dimensional simplicial complexes. This expansion is given by:

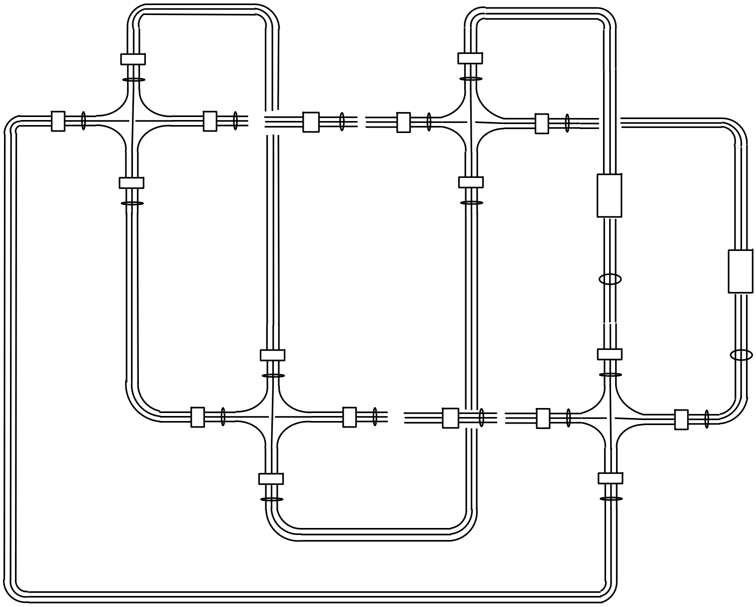
$$Z = \int \mathcal{D}\varphi \, e^{-S[\varphi]} = \sum_{\Gamma} \frac{\lambda^N}{sym[\Gamma]} \, Z(\Gamma),$$

where N is the number of interaction vertices in the Feynman graph Γ, $sym[\Gamma]$ is a symmetry factor for the graph, and $Z(\Gamma)$ the corresponding Feynman amplitude. We now sketch the calculation of the Feynman amplitudes of the model.

We first do so in the Lie algebra picture (following [12]), as this will immediately allow us to relate the amplitudes to simplicial gravity. We then use the propagator and vertex terms in (12.16). In building up the diagram, propagator and vertex strands are joined to one another using the $\star$-product. Each loop of strands bounds a face of the 2-complex Γ dual to a 3D simplicial complex Δ, bounded by several links dual to triangles sharing the same edge of Δ, in different tetrahedra. Because of the non-commutativity of the $\star$-product, it is important to choose and then keep track of the ordering between these triangles:



A rather involved, but still low-order, example of a Feynman diagram is represented in the following figure:



We first notice that, under the integration over group elements h_t, the amplitude factorizes into a product of face amplitudes. Let then f be a face of the 2-complex, and $A_f[h]$ its associated amplitude.

Choose an ordered sequence $\{\tau_j\}, \{t_j\}_{0\leq N}$, of triangles (to which propagators are associated) and tetrahedra (to which vertex functions are associated) surrounding the face (dual to an edge of the dual simplicial complex). Each propagator and each vertex function contributes a single "2-point function" $(\delta_{-x_i} \star \mathrm{e}_h)(y_i)$, to give:

$$A_f[h] = \int \prod_{j=1}^{N} \mathrm{d}x_j \vec{\bigstar}_{j=0}^{N+1} (\delta_{x_j} \star e_{h_{jj+1}})(x_{j+1}).$$

Next, we integrate over the N variables $x_1, \ldots, x_N$. Introducing the total holonomy $H_0 := h_{01} \cdots h_{N0}$ around the boundary of the face, the formula reduces to:

$$(e_{H_0} \star \delta_{H_0^{-1} x_0 H_0})(x_{N+1}) = (\delta_{x_0} \star e_{H_0})(x_{N+1}).$$

We now have to "close the loop" by setting $x_{N+1} = x_0$; in order to do so, we use $\delta_{x_0} = \int dg\, e_{g^{-1}}(x_0) e_g$, and get:

$$(\delta_{x_0} \star e_{H_0})(x_{N+1}) = \int dg\, e_{g^{-1}}(x_0) e_{gH_0}(x_{N+1}).$$

We see that we are left with a single integration over the Lie algebra, and we obtain:

$$A_f[h_t] = \int d^3x_0 \int dg \, e_{gH_0g^{-1}}(x_0).$$

Using $e_{gH_0g^{-1}}(x_0) = e_{H_0}(g^{-1}x_0g)$, a simple change of variable in x_0 leads to the Feynman amplitude:

$$Z(\Gamma) = \int \prod_t dh_t \prod_f A_f[h_t] = \int \prod_t dh_t \prod_f dx_f \, e^{i\sum_f \mathrm{Tr}\, x_f H_f}, \tag{12.20}$$

where H_f is the holonomy along the boundary of the face f, calculated for a given choice of a reference tetrahedron.

This is the usual expression for the simplicial path integral of 3D gravity in first-order form (or 3D BF theory). The Lie algebra variables x_f, one per edge of the simplicial complex, play the role of discrete triad, while the group elements h_t, one per triangle or link of the dual 2-complex, play the role of discrete connection, defining the discrete curvature H_f through holonomy around the faces dual to the edges of the simplicial complex [78, 81].

More precisely, consider the 3D gravity action in the continuum:

$$S(e,\omega) = \int_{\mathcal{M}} tr\,(e \wedge F(\omega)),$$

with variables the triad 1-form $e^i(x) \in \mathfrak{su}(2)$ and the 1-form connection $\omega^j(x) \in \mathfrak{su}(2)$, with curvature $F(\omega)$. Introducing the simplicial complex Δ and its topological dual cellular complex, we can discretize the triad in terms of Lie algebra elements associated with the edges of the simplicial complex as $x_e = x_f = \int_e e(x) = x^i\tau_i \in \mathfrak{su}(2)$, and the connection in terms of elementary parallel transports along links of Γ, dual to triangles of Δ, as $h_L = h_t = e^{\int_L \omega} \in \mathrm{SU}(2)$. The discrete curvature will then be given by the holonomy around the dual face f, obtained as an ordered product of group elements h_L associated with its boundary links: $H_f = H_e = \prod_{L\in\partial f} h_L = e^{F_f} \in \mathrm{SU}(2)$, and a discrete counterpart of the continuum action will be given by the action appearing in 12.20.

We have thus shown the Feynman amplitudes of the GFT model simplicial path integrals for 3D Riemannian gravity in first-order form. Moreover, in the non-commutative Lie algebra variables, the underlying simplicial geometry of the model is made transparent. The Lie algebra variables represent discrete triad variables associated with the edges of a triangle, in the corresponding frame, the GFT field being its second quantized wave function. The closure of the triangle is obtained by constraining appropriately these edge vectors. The GFT action encodes the correct gluing of triangles in a tetrahedron, and across tetrahedra, by identification of triad vectors up to parallel transport, parametrized by a gauge connection. The quantum dynamics of each interaction process of triangles, corresponding to a given simplicial complex/Feynman diagram, that emerges therefore as a virtual and quantum construction, is defined by the associated amplitude.

It is instructive to compute the GFT amplitude for given boundary simplicial data, i.e. for open GFT Feynman diagrams. The 1-vertex contribution to the 4-point functions is the function of 12 metric variables x_i, x_i' obtained by connecting a closure operator $\widehat{C}$ (propagator) on each external leg of the vertex diagram, thus building up four triangles $t_a, \ldots, t_c$. The resulting amplitude is the exponential of the BF action for a single (flat) simplex, and thus made out of pure boundary terms, with fixed metric (x variables) on the boundary (and integrated bulk connection). The vertex also relates the metric data of pairs of triangles t, t' via $(\delta_{x_i} \star \mathrm{e}_{h_{tt'}})(x_i')$. This may be viewed as a constraint on the gauge connection. In a semi-classical limit the constraints become those characterizing a discrete Levi–Civita connection.

This also implies that: first, for a generic simplicial complex with boundary, the GFT Feynman amplitudes in the x representation are given by a path integral for the BF action augmented by the appropriate boundary terms; second, the (exponential of the) BF action for a single simplex is already explicitly present in the interaction term of the GFT action, in the x representation. This can be useful to study the link with semi-classical/continuum gravity directly at the GFT level.

We have thus shown that the GFT model we have introduced succeeds in at least one of the points where the simpler tensor models failed, i.e. in defining amplitudes for its Feynman diagrams (identified with discrete spacetimes), arising in perturbative expansion around the "no-space state," that correctly encode the classical and quantum simplicial geometry and that can be related to a simplicial action for gravity.

As for the other problems that tensor models face, and that are also shared by GFTs, concerning the combinatorial structures obtained in the same perturbative expansion, and the control over this perturbative sum, we can hope that the additional structure and data in the GFT amplitudes, the same that allows the link with classical discrete gravity, would also help in progressing on this front. We will discuss in the following to what extent recent results either already fulfill or at least give more ground for this hope.

Notice also that it is well known that a field $\varphi \in \mathbb{C}(\mathrm{SU}(2)^3)$ can be seen as the tensor product of three representations of the quantum (Drinfeld) double $D\mathrm{SU}(2)$, which is a deformation of the Poincaré group [49, 62, 66]. This is the starting point of the analysis of the transformation properties of the Boulatov field, and of the corresponding symmetries of the GFT action we have presented above, and which have just recently been discovered [10]. The symmetries that one can identify at the level of the GFT action translate, at the level of the corresponding Feynman amplitudes, into the known symmetries of simplicial BF theory [45]. In particular, the so-called translation symmetry of BF theory (and 3D gravity), strictly related to diffeomorphism symmetry, has been identified in the GFT action, and related to (generalized) Bianchi identities in the Feynman amplitudes. This opens the door to the study of diffeomorphism symmetries at the GFT level, allowing then the application of field-theoretic techniques to the analysis of their consequences on GFT transition amplitudes.

The Feynman amplitudes can also be computed in the other representations. We first do so in the group picture, using the kinetic and vertex terms (12.18). The Feynman amplitudes are obtained by convolution of such vertex functions and propagators. The result of these convolutions is a single delta function $\delta(\prod_{L\in\partial f} h_L)$ for each 2-cell f in the Feynman diagram, dual to a single edge of the simplicial complex, with argument given by the product of group elements h_L each associated with a link in the boundary of the 2-cell. In other words, the model imposes flatness of the discrete curvature located on each dual 2-face, i.e. on each edge of the simplicial complex Δ. Once more, this is in accordance with our understanding of 3D simplicial gravity (and of continuum 3D gravity as well). The overall amplitude is then:

$$Z(\Gamma) = \prod_{L\in\Gamma} \int dh_L \prod_f \delta\Big(\prod_{L\in\partial f} h_L\Big)\,. \tag{12.21}$$

Of course, the same result could have been obtained by taking the group Fourier transform of the expression (12.20) for the same amplitude, which in this case would simply amount to performing the integral over x_f in (12.20).

Similarly, we can compute the spin foam expression of the Feynman amplitudes $Z(\Gamma)$, i.e. their expression in terms of group representations (quantum numbers of geometry). We can obtain it starting from the rewriting of the GFT action in representation space, i.e. from the expressions (12.20), or by direct Peter–Weyl decomposition of amplitudes (12.21), and successive group integrations. The result is an assignment of an irreducible SU(2) representation j_f to each face of Γ, and of a group intertwiner to each link of the same complex, i.e. a spin foam [72, 74, 81].

The corresponding amplitude reads:

$$Z(\Gamma) = \left(\prod_f \sum_{j_f}\right) \prod_f (2j_f + 1) \prod_v \left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{array} \right\}.$$

This is the well-known Ponzano–Regge spin foam model for 3D Riemannian quantum gravity, actually the first spin foam model ever proposed.

Many results have been obtained about this model, which, upon regularization, provides a topological invariant of 3D manifolds. We refer to the literature for details on such results, and a more detailed analysis of the model [19, 47, 48, 72, 74, 81]. For example, one can define a quantum group deformation of the Ponzano–Regge model, and of the corresponding GFT, passing from the group $SU(2)$ to the quantum group $SU_q(2)$ and to the corresponding category of representations, and obtain in this way the so-called Turaev–Viro model, another invariant of 3-manifolds [74, 91], and related to 3D gravity with positive cosmological constant.

One can show, for a single spin foam vertex (tetrahedron), that, for large, j's, which can be understood as a semi-classical approximation of the amplitude:

$$\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{array} \right\}_{v*} \simeq \cos S_R(l_e) \simeq e^{iS_R} + e^{-iS_R},$$

where $S_R(l_e = 2j+1)$ is the Regge action for simplicial gravity, with edge lengths given by $2j_e + 1$. Thus the Feynman amplitudes in spin foam representation match the expected form of semi-classical simplicial gravity path integral. This is not surprising, given the dual simplicial path integral expression for the same amplitudes in terms of Lie algebra variables x_e, at least from a heuristic point of view. The x_e in fact are interpreted as classical edge vectors (discrete triads), whose absolute value is the length L_e of the edge to which they are associated, and that has instead quantum spectrum (according to the canonical theory) $L_e = \sqrt{j_e(j_e+1)} \approx 2j_e + 1$, where the approximation holds true for large j_e (semi-classical limit). In this approximation, in fact, the quantum numbers (eigenvalues) are expected to match the corresponding classical variables, and thus also the amplitude should have a functional dependence on them that matches the dependence on the triad variables, in the same approximation (large scales). Now, look at the simplicial path integral we obtained in the Lie algebra representation, for a single tetrahedron and for fixed x variables on the boundary. Its semi-classical limit implies a saddle-point approximation of the action (as in any path integral), which, together with the (commutative limit of) the delta functions on the algebra entering the amplitude, impose that the discrete connection is a Levi–Civita one and a function of the triad variables x, thus resulting in a second-order action in terms of them only, i.e. in the Regge action for simplicial gravity. Seen in this light, the asymptotic result for the $6j$-symbol only confirms the geometric meaning of the quantum variables j_e and would suggests in itself (if we didn't have it already) a simplicial gravity path integral formulation of the same spin foam amplitudes.

We have thus shown that the Feynman amplitudes of the GFT model can equivalently be written in the form of a spin foam model and a simplicial gravity path integral. This is an exact duality that stems from the possibility of using two equivalent representations of the GFT field, represented as a function on group manifolds: as a function of representation labels, following Peter–Weyl decomposition, and as a non-commutative function on Lie algebra variables, using the non-commutative group Fourier transform.

Finally, before moving on to the 4-dimensional case, let us comment on the calculation of transition amplitudes. Transition amplitudes or correlation functions between quantum gravity states are defined, as insertions of appropriate GFT observables in the partition function of the theory. These observables are, as in any field theory, arbitrary functionals of the fundamental GFT field, compatible with the symmetries of the theory. Polynomial functionals, in particular, can be used as a basis for the space of observables, and are obtained by convolution in the Lie algebra or in the group variables of products of GFT fields, or by the equivalent traces in representation space. As we hinted at above, such observables describe extended simplicial spaces or, equivalently, extended spin networks, endowed with quantum geometric data, and thus define possible quantum spaces. When inserted in the GFT path integrals, and upon expansion of the same in Feynman diagrams, they provide the boundary structures for the open diagrams, and the perturbative sum provides a tentative definition of their mutual correlations (transition amplitudes) [39, 79].

For example, one can consider the spin network observables

$$O_{\Psi=(\gamma,j_e,i_v)}(\phi) = \left(\prod_{(ij)\int} dg_{ij}dg_{ji}\right)\Psi_{(\gamma,j_e,i_v)}(g_{ij}g_{ji}^{-1})\prod_i \varphi(g_{ij}),$$

where $\Psi_{(\gamma,j_e,i_v)}(g)$ identifies a spin network functional [85, 90] for the spin network labeled by a graph γ with representations j_e associated with its edges and intertwiners i_v associated with its vertices, and g_{ij} are group elements associated with the edges (ij) of γ that meet at the vertex i. The transition amplitude between boundary data represented by two spin networks, of arbitrary combinatorial complexity, can be expressed as the expectation value of the field operators having the same combinatorial structure as the two spin networks [39].

$$\langle \Psi_1 \mid \Psi_2 \rangle = \int \mathcal{D}\varphi\, O_{\Psi_1}\, O_{\Psi_2}\, e^{-S(\varphi)} = \sum_{\Gamma/\partial\Gamma=\gamma_{\Psi_1}\cup\gamma_{\Psi_2}} \frac{\lambda^N}{sym[\Gamma]}\, Z(\Gamma),$$

where the sum involves only 2-complexes (spin foams) with boundary given by the two spin networks chosen.

12.3 Towards a group field theory formulation of 4D quantum gravity

Important results have been obtained recently in the attempt to construct interesting and, possibly, complete GFT and spin foam models for 4-dimensional quantum gravity. Here, we summarize some of them briefly.

The Boulatov model can be generalized easily to any dimension (and any compact group), and one can similarly generalize both the group Fourier transform and the Peter–Weyl decomposition, to write GFT field, action and amplitudes in Lie algebra, group or representation variables. In all cases, it defines a GFT quantization of BF theory in the given dimension and for the given gauge group. In particular, the 4-dimensional extension of the Boulatov model, proposed by Ooguri [71] is based on a real field: $\varphi(g_1,\ldots,g_4):\mathrm{SO}(4)^{\times 4}\to\mathbb{R}$, symmetric under diagonal left action: $\varphi(hg_1,hg_2,hg_3,hg_4)=\varphi(g_1,g_2,g_3,g_4)$. Equivalently, it can be expressed as a field over four copies of the Lie algebra $\mathfrak{so}(4)$ with the four Lie algebra arguments satisfying the closure condition $x_1+x_2+x_3+x_4=0$.

The GFT action is then:

$$\begin{aligned} S[\phi] = &\frac{1}{2}\int [\phi(g_1,g_2,g_3,g_4)]^2 \\ &-\frac{\lambda}{5!}\int \phi(g_1,g_2,g_3,g_4)\phi(g_4,g_5,g_6,g_7)\phi(g_7,g_3,g_8,g_9) \\ &\times\phi(g_9,g_6,g_2,g_{10})\phi(g_{10},g_8,g_5,g_1)\,, \end{aligned}$$

or the equivalent in Lie algebra variables (with appropriate $\star$-multiplications).

The combinatorics now represents the gluing of five tetrahedra (each corresponding to a field φ) pairwise along boundary triangles (each corresponding to one of the four arguments of φ) to form a 4-simplex. The Feynman diagrams are 2-complexes dual to 4D simplicial complexes.

Kinetic and vertex terms again encode the identification up to parallel transport of the bivectors associated with the same triangle in different tetrahedral frames, and the computation of Feynman amplitudes proceeds analogously to the 3D case. The result is again a simplicial path integral for BF theory like (12.20), with integrals now over SO(4) group and Lie algebra elements [12].

The Lie algebra variables are interpreted as the four bivector variables associated with the four triangles of each tetrahedron in a simplicial discretization of 4D BF theory, coming from the discretization of the continuum B field [29, 83]. The continuum 1-form connection is instead discretized as in the 3D case, in terms of group elements associated with links of the 2-complex dual to the simplicial complex, to give holonomies (discrete curvature) associated with each 2-cell to a triangle of the same.

Similarly, one can obtain a spin foam expression for the same Feynman amplitude. It is given by:

$$Z(\Gamma)=\prod_f\int_{\mathfrak{so}(4)} dx_f \prod_L \int_{\mathrm{SO}(4)} dh_L\, e^{i\sum_f tr(x_f H_f)}$$

$$= \left(\prod_{L \in \Gamma} \int dh_L \right) \prod_f \delta \left(\prod_{L \in \partial f} h_L \right)$$
$$= \sum_{\{j_+, j_-\}} \prod_f (2j_+ + 1)(2j_- + 1) \prod_v \{15 - j\}_+^v \{15 - j\}_-^v, \qquad (12.22)$$

where we used the self-dual/anti-self-dual splitting of SO(4) representations, and the symbol for the vertex amplitude is the well-known $15j$-symbol from the recoupling theory of angular momentum. We also know how the spin foam vertex amplitudes can be reconstructed from the structure of the spin network boundary states, with SO(4) labels. As in 3D, one considers one SO(4) intertwiner for each tetrahedron in the boundary of a 4-simplex, with one representation associated with each triangle of the tetrahedron, and then glues these intertwiners along common magnetic indices, following the combinatorial pattern of triangles in the boundary of a 4-simplex.

12.3.1 Gravity as a constrained topological BF theory

Now this easy generalization of the GFT and spin foam formalism from 3D to 4D BF theory would not be very important if not for a simple fact: classical 4D gravity can be expressed as a constrained BF theory (Plebanski formulation) [29, 83]. For an $\mathfrak{so}(4)$ Lie algebra 1-form connection ω and a 2-form B also with values in the Lie algebra, one can write the classical continuum action:

$$S(\omega, B, \phi) = \int_{\mathcal{M}} \left[B^{IJ} \wedge F_{IJ}(\omega) - \frac{1}{2} \phi_{IJKL} B^{KL} \wedge B^{IJ} \right],$$

with the obvious symmetries for the Lagrange multipliers $\phi_{[IJ][KL]} = -\phi_{[IJ][KL]}$. Variations with respect to these Lagrange multipliers ϕ give constraints on the B variables, whose solution (in the appropriate sector [74, 81]) forces the same B to be a function of a tetrad field e as: $B^{IJ} = \pm \frac{1}{2} \varepsilon^{IJ}{}_{KL} e^K \wedge e^L$, so that, for these solutions, the action becomes the Palatini action for gravity:

$$S(\omega, e) = \int_{\mathcal{M}} \left[\frac{1}{2} \varepsilon^{IJ}{}_{KL} e^K \wedge e^L \wedge F_{IJ}(\omega) \right].$$

As we have already seen, we know well how to discretize and quantize BF theories in any dimension, and in particular we know how to construct the corresponding spin foam models and GFT action. Furthermore, we also know how to discretize the Plebanski constraint on an arbitrary simplicial complex [1, 16, 29, 37, 42, 83]. In fact, a discrete tetrad can be reconstructed for the whole simplicial complex, and it determines the discrete bivectors B_f associated with triangles of the same, if one requires that

$$\forall \text{tetrahedra}\, t \in \Delta \;\; \exists n_t \in S^3 \;/\; (*B_f)^{IJ} n_{tJ} = 0 \;\; \forall B_f \;\; f \subset t,$$

i.e. if one requires that, for each tetrahedron, its four triangle bivectors all lie in the same hypersurface orthogonal to the same normal vector, interpreted as normal to the tetrahedron, together with closure of all the tetrahedra in Δ, and an orientation and a non-degeneracy condition [16, 29, 83].

All this can be generalized to include an Immirzi parameter γ (which plays a crucial role in LQG), adding a term to the Plebanski action:

$$S(\omega, B, \phi) = \int_{\mathcal{M}} \left[B^{IJ} \wedge F_{IJ}(\omega) + \frac{1}{\gamma} B^{IJ} \wedge F_{IJ}(\omega) - \frac{1}{2} \phi_{IJKL} B^{KL} \wedge B^{IJ} \right].$$

On the solution to the constraints, this becomes the Holst action for gravity:

$$S(\omega, e) = \int_{\mathcal{M}} \left[\frac{1}{2} \varepsilon^{IJ}{}_{KL} e^K \wedge e^L \wedge F_{IJ}(\omega) + \frac{1}{2\gamma} e^I \wedge e^J \wedge F_{IJ}(\omega) \right],$$

which is classically equivalent to Palatini gravity, as the additional term vanishes on-shell. This action is the classical starting point for the loop quantization leading to LQG.

Also, the discrete theory can easily be generalized to include the Immirzi parameter. One has only to impose on the BF action the new constraints:

$$\forall \text{tetrahedra}\, t \in \Delta \;\; \exists n_t \in S^3 \;/\; (B_f - \gamma * B_f)^{IJ} n_{tJ} = 0 \;\; \forall B_f \;\; f \subset t.$$

12.3.2 A basic idea and two strategies for 4D gravity model-building in GFT

A natural procedure[3] for defining a GFT or spin foam model for 4D gravity, given the knowledge we have on BF models, is then to start from the GFT action for 4D BF theory, or from the corresponding spin foam expression, and impose suitable restrictions on the dynamical variables representing the discrete (classical or quantum) B-field variables and so to impose the discrete version of the Plebanski constraints. The result, if things are done properly, should be a good encoding of 4-simplicial (and possibly continuum) geometry.

On the basis of the duality between simplicial gravity path integrals and spin foam models, naturally realized in a GFT context, as we have seen, we can follow two possible strategies:

[3] But certainly not the only one [8].

1. Find a quantum version of Plebanski constraints, impose them directly at the quantum level on BF spin network states to get gravity spin network states, then construct spin foam/GFT amplitudes from these states; finally, check that they encode correctly simplicial geometry by relating them to a simplicial gravity path integral.
2. Start from the BF GFT model in terms of Lie algebra variables (bivectors) and impose the geometric constraints on them, obtaining a model whose Feynman amplitudes are manifestly simplicial path integrals for BF with gravity constraints being imposed; then rewrite them as spin foam models and identify the corresponding space of boundary states.

Let us discuss these two strategies in turn and report briefly on the results obtained recently by following them.

The state sum strategy and spin foam models

The first strategy is based on geometric quantization of simplicial structures or on the use of master constraint techniques, and is the traditional and most developed one [16, 35, 37]. The idea is the following. The classical B variables, at the quantum level, are identified with generators $T^{IJ} = (T_+^i; T_-^j)$ using the usual self-dual/antiselfdual decomposition of the $\mathfrak{so}(4)$ Lie algebra, which then act as operators on the SO(4) spin network states labeled by representation of the same group. The classical constraints, functions of the B's, are then also turned into operators and imposed at the quantum level on such spin networks (but up to operator ordering corrections), to give restrictions on both the representations attached to their links and on the intertwiners associated with their vertices [37]. The resulting spin networks satisfying these restrictions are identified as *gravity spin networks*, built out of *gravity intertwiners* and used to construct new spin foam vertex amplitudes. This construction follows the analogous one for the BF vertex amplitude: take five gravity spin network vertices (corresponding to five tetrahedra in the boundary of a 4-simplex); "glue them together" by tracing out internal (vector) variables in each common representation space (corresponding to triangles common to two tetrahedra). This results in the spin foam amplitude for a single 4-simplex.

An alternative, but closely related way of imposing the same constraints [35–37, 42, 67] makes use of the Perelomov coherent states for groups [80]. It amounts to the following: re-write the BF spin foam in terms of SO(4) coherent states $|j^+, j^-, (n^+, n^-)\rangle$, where $n^\pm \in S^2$; identify the coherent state parameters (n^+, n^-) as semi-classical counterparts of classical B variables, and impose the gravity constraints on them. This is sensible because the number of components matches, and their behavior under SO(4) rotations. Also their semi-classical behavior matches.

In fact,

$$\langle j^+, j^-, (n^+, n^-)|(T^i_+, T^i_-)|j^+, j^-, (n^+, n^-)\rangle = (j^+ n^i_+, j^- n^i_-)\,.$$

These constraints on coherent state parameters translate into constraints on the representations of SO(4), and thus on the quantum states allowed in the model.

In order to appreciate the differences between the various models that have been constructed using this strategy, it is convenient to use the decomposition of each representation (j^+, j^-) of SO(4) into irreducible representations of the diagonal SU(2) subgroup k. There are 10 representations, one per triangle in the 4-simplex, each triangle being shared by two tetrahedra. The amplitudes all depend on 10 representations of SO(4) and on the 10 representations of SU(2), and on the corresponding intertwiners, one for each of the five tetrahedra in the 4-simplex, ensuring SO(4) gauge invariance. The general form of the vertex (4-simplex) amplitude is:

$$A_v(k_f, i_e) = \sum_{j^+_{ab}, j^-_{ab}} \{15j^+\}\{15j^-\} \prod_{(ab)=1,..,10} f(j^-_{ab} j^+_{ab}, k_{ab}) \prod_{a=1,..,5} i_a(k_{ab}),$$

where the indices a, b label the five tetrahedra and the pairs (ab) label triangles. The functions f are appropriate mapping coefficients from SO(4) to SU(2) representations and they differentiate between the different models.

The resulting models, for various choices of γ, are [35–37, 42, 67]:

- In the geometric sector ($\gamma \to \infty$), one obtains the Barrett–Crane vertex, characterized by: $j^+_{ab} = j^-_{ab}$, $k_{ab} = 0$; using the coherent state method, one instead obtains the Freidel–Krasnov vertex, with a more involved dependence from k_{ab} on.
- In the topological sector ($\gamma \to 0$), one obtains the Engle–Pereira–Rovelli vertex, with $k_{ab} = 2j^+_{ab} = 2j^-_{ab}$.
- In the case of finite Immirzi parameter, one obtains the Engle–Pereira–Rovelli–Livine vertex, with $j^+_{ab} = \frac{\gamma+1}{|\gamma-1|} j^-_{ab}$ and $k_{ab} = j^+ \pm j^-_{ab}$ in the cases $\gamma < 1$, and $\gamma > 1$, respectively; using the coherent state method one obtains the more complicated Freidel–Krasnov model with Immirzi parameter, which, however, for $\gamma < 1$ coincides with the Engle [37] one.

To conclude our brief survey of the proposals on the table, we notice the good and less good points of this first strategy. The method provides nicely the structure of boundary states, and a reasonably simple vertex amplitude compatible with the symmetries (gauge invariance) and with the gravity constraints (at least in a semi-classical approximation). Thanks to the use of master constraint techniques [90], it nicely takes care of constraint classes, and so imposes different classes of constraints appropriately. On the negative side, it does not specify uniquely the other contributions to the spin foam amplitudes, i.e. triangle and tetrahedral amplitudes, and does not give full control on the quantum regime, i.e. beyond the semi-classical regime in which the solutions to the quantum constraints have been identified.

Moreover, these models, currently under active investigation, have already been shown to have two very nice properties: (1) their boundary spin networks are of the same type as those obtained in canonical loop quantum gravity, and also the kinematical geometric operators such as areas and volumes have the same spectrum as that obtained in loop quantum gravity – this is a very important matching between the canonical and the covariant construction; (2) the vertex amplitudes reduce to the (cosine of) the Regge action in a semi-classical (large-j) limit, confirming that the models correctly capture simplicial geometry at least for a single 4-simplex and at least in such an approximation [17, 19, 26].

Despite these successes, and leaving aside technical issues in their derivation, the construction of the new models is still half of the story, from the GFT point of view. They can be given a GFT formulation [42], in representation space, but we have no clear understanding of the geometric translation of the constraints on representations in terms of group (connection) variables for $\gamma \to \infty$. We lack, therefore, a clear group picture for the corresponding GFT. We miss as well a reformulation in terms of Lie algebra elements. As a consequence, we do not have under full control the geometric meaning of the new models, nor a simplicial path integral representation of the same, in terms of a constrained BF with Lie algebra and group variables, with a definite choice of measure (but see [27]).

The (non-commutative) geometric strategy and simplicial path integrals

Let us now pass to the second possible strategy, which starts from the Lie algebra representation of the Ooguri model, which provides in fact a convenient starting point for imposing in a geometrically transparent manner the discrete gravity constraints. Notice that simplicial BF in 4D (and Euclidean signature) has a classical phase space, which decomposes into $T^*\mathrm{SO}(4)$ for each triangle of the simplicial complex (similarly for the continuum theory, quantized as LQG). This justifies the use of the SO(4) non-commutative group Fourier transform. This line of research is, however, less developed, having started only very recently, when the non-commutative Lie algebra representation for GFT was developed [12]. We report briefly on its current status.

One can easily impose that the four bivectors in each tetrahedron are orthogonal to the same normal vector $n \in \mathcal{S}^3 \simeq \mathrm{SU}(2)$ to the tetrahedron, by means of the constraint:

$$\widehat{S}_n(x_j^-, x_j^+) = \prod_{j=1}^{4} \delta_{-nx_j^- n^{-1}}(x_j^+),$$

where we have used the decomposition of the $\mathfrak{so}(4)$ algebra into self-dual and anti-self-dual components, and we indicate with $\widehat{.}$ functions in the Lie algebra space. Using SO(3) delta functions, one obtains a constrained field:

$$\begin{aligned}(\widehat{S}_n \star \widehat{\varphi})(x) &= \int_{\mathrm{SO}(4)} dg\, \varphi(g) \int_{\mathrm{SO}(3)} du\, e_{n^{-1}ung^-}(x^-) e_{ug^+}(x^+) \\ &= \int_{\mathrm{SO}(4)} dg\, S_n\varphi(g)\, e_{g^-}(x^-) e_{g^+}(x^+),\end{aligned}$$

where

$$S_n\varphi(g) := \int_{\mathrm{SO}(3)} \varphi(n^{-1}ung^-, ug^+),$$

having decomposed the SO(4) element g into a pair of SO(3) elements $g^{\pm}$ following the splitting of the covering group $Spin(4) \sim \mathrm{SU}(2) \times \mathrm{SU}(2)$. Hence, $\widehat{S}_n$ is the dual of the projector onto the fields on a homogeneous space $\mathcal{S}^3 \sim \mathrm{SO}(4)\backslash\mathrm{SO}(3)_n$ obtained as a quotient of SO(4) with the subgroup $\mathrm{SO}(3)_n$ of elements which stabilize the normal n. The standard Barrett–Crane projector [29] is obtained for $n = 1$, i.e. for the diagonal SO(3) subgroup.

By combining the simplicity projector $\widehat{S} := \widehat{S}_1$ with the closure projector one gets the field $\Psi(x) \equiv (\widehat{S} \star \widehat{C} \star \widehat{\varphi})(x)$, which can be shown to reproduce the GFT field used in the standard GFT formulation of the Barrett–Crane model. More precisely, combining the interaction term:

$$\frac{\lambda}{5!} \int \Psi_{1234}\Psi_{4567}\Psi_{7389}\Psi_{96210}\,\Psi_{10851}$$

with the possible kinetic terms:

$$\frac{1}{2}\int \widehat{\varphi}_{1234}\widehat{\varphi}_{4321} \text{ or } \frac{1}{2}\int \Psi_{1234}\Psi_{4321}\,,\ \frac{1}{2}\int (\widehat{C} \star \widehat{\varphi})_{1234}(\widehat{C} \star \widehat{\varphi})_{4321}$$

one obtains, respectively, the versions of the BC model derived in [23, 30, 82]. The origin of these different versions can be understood geometrically, in the Lie algebra representation of the GFT. For $h \in \mathrm{SO}(4)$, we have:

$$(e_h \star \widehat{S}_n)(x) = (\delta_{-nh_-^{-1} \cdot_- h_- n^{-1}}(h_+^{-1} \cdot_+ h_+) \star e_h)(x) = (\widehat{S}_{h \rhd n} \star e_h)(x),$$

where $h \rhd n := h_+ n h_-^{-1}$. This expresses the fact that, after rotation by h, simple bivectors with respect to the normal n become simple with respect to the rotated normal $h\triangleright$. So, closure and simplicity constraints do not commute (see also [73]). Geometrically it means that the normal variables n are not transformed alongside the bivectors under SO(4) frame rotations, as they should.

A simplicial path integral formulation of the Barrett–Crane model, for, say, the BC version [23], is obtained by using the propagator and vertex:

$$\prod_{i=1}^{4}(\delta_{-x_i^{\pm}})(y_i^{\pm}), \quad \int \prod_{t} dh_{\tau\sigma} \prod_{i=1}^{10}(\delta_{-x_i^{\pm}} \star \widehat{S} \star e_{h_{\tau\tau'}})(y_i^{\pm}),$$

leading to an integral over connection variables $h_{\tau\sigma}$ and Lie algebra elements x_f of the product of face amplitudes

$$\begin{aligned} A_f &= \int dg\, (e_{g^{-1}} \star \{\delta_{-\bullet^-}(\bullet^+) \star e_{h_{01}} \star \cdots \delta_{-\bullet^-}(\bullet^+) \star e_{h_{N0}}\} \star e_g)(x_f) \\ &= \int dg\, (\mathcal{O}_f \star e_H)(g x_f g^{-1}) = A_f(x_f, h_{\tau\sigma}), \end{aligned}$$

where the observable $\mathcal{O}$ is a $\star$-product of $N+1$ simplicity constraints written in the frame of the $N+1$ tetrahedra, and suitably parallel-transported:

$$\mathcal{O}_f(x) = \bigstar_{j=0}^{N} \delta_{-h_{0j}^{-\,-1} x^- h_{0j}^-}(h_{0j}^{+\,-1} x^+ h_{0j}^+).$$

The integers $j = 0, \ldots, N$ label an ordered sequence of tetrahedra around the face f; x_f is the metric variable attached to f in the frame of the reference tetrahedron 0.

Finally, the whole Feynman amplitude takes the form of a non-commutative observable insertion in BF theory:

$$Z_{BC} := \int \prod_{\tau\sigma} dh_{\tau\sigma} \int \prod_{f} d^3x_f\, (\mathcal{O}_f \star e_{H_f})(x_f).$$

Therefore, we see that the Lie algebra representation of the GFT model allows: (1) a geometrically transparent imposition of the constraints on the x bivector variables; (2) some dubious aspects in the standard GFT derivations of the BC model to be identified; (3) a simplicial path integral rewriting of the BC spin foam amplitudes.

Now, the next step along this line of research would be: (1) to modify the simplicity projector, to obtain models for the topological sector and for the case of finite Immirzi parameter; (2) to look for a modification of the whole construction that solves the problematic geometrical issues we have identified in the standard BC GFTs.

Work on this is in progress [13], and can either lead to the definition (geometrically well-motivated) of different foam models for 4D quantum gravity, or to a complete and geometrically clear GFT formulation of the recently proposed ones, and, in one stroke, to a reformulation of the same as simplicial path integrals.

12.4 A selection of research directions and recent results

We now present a selection of recent results, obviously reflecting (and limited by) our own taste and knowledge.

12.4.1 Making sense of the quantum field theory: combinatorics of Feynman diagrams and GFT renormalization

A first area of recent developments [20, 21, 40, 58, 68] has been the application of quantum field theory techniques to GFTs, to gain a better understanding and control over its perturbative expansion, using tools from renormalization theory, which is and will be relevant both for a mathematical definition of the theory and for the study of its continuum approximation. The aim is thus to address the remaining main shortcomings of tensor models, within the GFT approach.

As we said, GFTs define a sum over simplicial complexes (1) of arbitrary topology and (2) that correspond, in general, to pseudo-manifolds, (the dual d-cells have boundaries which are not $(d-1)$-spheres). The issue of controlling the sum over topologies, and of identifying an approximation in which simple topologies dominate, has an analogue in the context of matrix models [28, 53], and it was solved by the definition of the large-N limit, in which diagrams of trivial topology (S^2 in the compact case) dominate the Feynman amplitudes of the theory. One goal, in this respect, in the GFT context, would be to define some generalized version of this scaling limit. In turn, this requires understanding in detail the degree of divergence of arbitrary GFT diagrams, and possibly characterizing it in terms of topological quantities. Notice that the issue of controlling the relative weight of manifolds and pseudo-manifolds in the perturbative sum, and possibly identifying a regime in which the second are subdominant, arises instead only in dimensions $d > 2$. Notice also that these issues, which are basically the issues of divergence of spin foam amplitudes, can now be tackled using standard field theory language and techniques, in the GFT framework; more precisely, they become a problem in GFT renormalization.

The work of [40] took the first steps toward solving these three issues, starting a systematic study of GFT renormalization, in the context of the Boulatov model for 3D (Riemannian) quantum gravity (but the results apply to a wider class of models). We have seen that the Feynman diagrams of the theory are, by construction, 3D triangulations, while the corresponding Feynman amplitudes are given by the

Ponzano–Regge spin foam model [19, 47, 74], which provides a quantization of 3D gravity discretized on the simplicial complex dual to the given Feynman diagram.

The divergences of this model are related to the topology of the bubbles (3-dimensional cells), dual to vertices of the simplicial complex, in the Feynman diagrams, but it is difficult to establish which diagrams need renormalization in full generality, mainly due to the very complicated topological structure of 3D simplicial complexes, after a scale is introduced in the theory by an explicit cut-off in the spectral decomposition of the propagator.

What is achieved in [40] is:

- A detailed algorithm is given for identifying bubbles (3-cells) in the Feynman diagrams of the model, together with their boundary triangulations, which in turn can be used to identify the topology (genus) of the same boundary.
- Using this algorithm, the identification a subclass of Feynman diagrams which allow for a complete contraction procedure, and thus the ones that allow for an almost standard renormalization; this class of graphs, dubbed "type 1," is a natural generalization of the 2D planar graphs of matrix models (in the sense that the 2D restriction of their definition identifies indeed planar diagrams), thus suggesting that they can play a similar role in GFTs to that of planar diagrams in matrix models.
- For this class of diagrams, an exact power counting of divergences is proven, according to which their divergence is of the order:

  $$A_\Gamma = \left(\delta^\Lambda(I)\right)^{|\mathcal{B}_\Gamma|-1},$$

  where $|\mathcal{B}_\Gamma|$ is the number of bubbles in the diagram Γ, and $\delta^\Lambda(I)$ is the delta function on the group, with cut-off Λ, evaluated at the identity I.

A different perspective on divergences in GFT amplitudes is taken in [68], which also tackles the difficult issue of the summability of the entire perturbative sum (thus including the sum over topologies). The authors consider both the Boulatov model and a modification of the same proposed in [45]. The modification amounts to adding a second interaction term in the action, given by:

$$+\frac{\lambda\,\delta}{4!}\prod_{i=1}^{6}\int dg_i\,[\phi(g_1,g_2,g_3)\phi(g_3,g_4,g_5)\phi(g_4,g_2,g_6)\phi(g_6,g_5,g_1)]. \quad (12.23)$$

The new term corresponds simply to a slightly different recoupling of the group/representation variables, geometrically corresponding to the only other possible way of gluing four triangles to form a closed surface. This modification gives rise to a Borel summable partition function [45], which shows that a control over the sum of topologies and a non-perturbative definition of the corresponding GFT is feasible.

For both the Boulatov model and the modified one, the authors of [68] manage to establish very general perturbative bounds on amplitudes using powerful and elegant constructive techniques, rather than focusing on explicit power counting or Feynman evaluations, or on the combinatorial structures of the diagrams. They find that, the amplitudes of the Boulatov model for a diagram with n vertices are bounded, with cut-off Λ, by $K^n \Lambda^{6+3n/2}$, for some arbitrary positive constant K, while the modifed model has amplitudes bounded by $K^n \Lambda^{6+3n}$. Both bounds can be saturated. In fact, those that saturate the bound in the Boulatov case are type 1 graphs in the definition of [40].

The second main result of [68] relies again on constructive field theory techniques. A cactus expansion of the BFL model is obtained, and used to prove the Borel summability of the free energy of the model and to define its Borel sum. We can expect more interesting applications of these techniques to other GFT models, also in higher dimensions, in the future.

A variation of the above group field theory, in any dimension, has been introduced in [57] and it has already been proven to be very useful in the topological analysis of the GFT Feynman diagrams, in GFT renormalization, and in the study of GFT symmetries.

The colored version of the 3D gravity GFT model we have presented. This is defined in terms of four complex fields over $\mathrm{SU}(2)^3$: $\varphi_f(g_1, g_2, g_3)$, $f = 1, 2, 3, 4$, with the same diagonal invariance (or closure condition), by the action:

$$
\begin{aligned}
S_{3d}[\phi] = &\ \frac{1}{2}\sum_t \int [dg]^3 \varphi_t^*(g_1, g_2, g_3)\varphi_t(g_3, g_2, g_1) \\
&- \frac{\lambda}{4!}\int [dg]^6 \varphi_1(g_1, g_2, g_3)\varphi_2(g_3, g_4, g_5)\varphi_3(g_5, g_2, g_6)\varphi_4(g_6, g_4, g_1) \\
&- \frac{\lambda^*}{4!}\int [dg]^6 \varphi_1^*(g_1, g_2, g_3)\varphi_2^*(g_3, g_4, g_5)\varphi_3^*(g_5, g_2, g_6)\varphi_4^*(g_6, g_4, g_1),
\end{aligned}
$$

where $*$ denotes complex conjugation.

This model corresponds to having "colored" the four triangles t of each tetrahedron, and having imposed that only triangles of the same color can be identified when gluing tetrahedra together in the Feynman expansion. Because of this restriction, and making use of the complex structure to define an orientation in the Feynman diagrams, it turns out that only orientable complexes are generated by the perturbative expansion, regardless of the ordering of arguments chosen in each field.

The colored models can be seen as the GFT equivalent of the multi-matrix models, [28, 53].

These colored GFT models (in any dimension) have several other nice features.

- First of all, the possibility of giving a clear definition of bubbles, i.e. of the 3-cells of the Feynman diagrams of any dimension, which means that the colored Feynman diagrams identify a full d-dimensional cellular complex, and not just a 2-complex, i.e. a spin foam.
- Second, the possibility of defining a (not standard) computable cellular homology for each Feynman diagram. This can be used to show that the dual simplicial complexes are manifolds or pseudo-manifolds, as usual, but with only point-like singularities allowed in the non-manifold case [38].
- Moreover, one can prove the absence of generalized "tadpoles" and of "tadfaces" [20]. All this simplifies greatly the analysis of divergences in these models.
- Indeed, the authors of [20] can improve considerably, for colored models, the scaling bounds obtained in [68] for non-colored models.
- If the field φ is assumed to be fermionic, the model enjoys an SU(4) global symmetry in the field indices which is, however, absent in the bosonic case.
- This "coloring" turns out to be crucial also for the analysis of GFT symmetries, and in particular for identifying the GFT counterpart of translation symmetry of BF theories, strictly related to discrete diffeomorphism symmetry [9].
- One can also define a related homotopy transformation and show the link between the GFT amplitudes to the fundamental group of the cellular complex.
- One can prove that the corresponding closed "type 1 graphs" are homotopically trivial, and thus, by the Poincaré conjecture, 3-spheres. This proves the first conjecture put forward in [40].

Moreover, in [58], the boundary graphs of the same colored models are identified, and the topological (Bollobas–Riordan) Tutte polynomials associated with (ribbon) graphs are generalized to topological polynomials adapted to colored group field theory diagrams in arbitrary dimension.

A last development in the analysis of GFT renormalization, for BF models in arbitrary dimension, was reported in [21]. The authors study a version of the BF GFTs obtained by replacing, everywhere in the GFT Feynman amplitudes, the group G of dimension d with the additive commutative group of a vector space $\mathbb{R}^d$. This can be seen as the result of a Taylor expansion around the unity of the group for each group element entering the Feynman amplitudes, and retaining only linear terms in the Lie algebra elements. For this class of models, [21] establishes the exact power counting of arbitrary Feynman diagrams in terms of a certain incidence matrix between lines and faces. In the colored GFT case, one can do more and relate the same power counting to the homology of the Feynman diagrams, as has been defined in [57]. More precisely, one obtains, for an arbitrary connected vacuum diagram, the following degree of divergence in the cut-off Λ:

$$Z(\Gamma) \simeq K\, \Lambda^{\sum_{k=3}^{d+1}(-1)^{k-1}|\mathcal{B}_k| + \sum_{k=2}^{d-1}(-1)^k h_k},$$

where K is an arbitrary constant, $|\mathcal{B}_k|$ the number of k-bubbles of the Feynman diagram, with $|\mathcal{B}_{d+1}| \equiv 1$, and h_k is the kth Betti number, i.e. the dimension of the kth homology H_k of the diagram.

In particular, for type 1 graphs in $d = 3$, one finds that the degree of divergence is: $|\mathcal{B}_3| - 1 + h_2$, so that they must have $h_2 = 0$.

We refer the reader to the cited literature for the analysis of the properties of these models, and for the proofs of the mentioned results.

The authors of [21] also motivate their linearized models, arguing that they can be seen as an approximation of the full models in the limit of large representations. Work is in progress to apply the non-commutative GFT representation of [12], in terms of Lie algebra variables, to understand in more detail if and how the general models reduce to the linearized ones, and thus to check to what extent the degree of divergence found for them holds true in the general case [11].

Extracting physics: deriving non-commutative effective matter field theories from GFT

The other set of results we want to mention are interesting steps in the direction of bridging the gap between the microscopic description of quantum space, as provided by the GFT perturbative expansion (and the language of spin networks, simplices, spin foams, etc.) and macroscopic continuum physics, including usual general relativity and quantum field theories for matter. This is the outstanding problem faced by *all* current discrete approaches to quantum gravity and by loop quantum gravity as well, in spite of its continuum nature [78].

One would expect [75] a generic continuum spacetime to be formed by zillions of Planck-size building blocks, rather than few macroscopic ones, and thus to be, from the GFT point of view, a many-particle system whose microscopic theory is given by some fundamental GFT action. This also suggests looking for ideas and techniques from statistical field theory and condensed matter theory, and trying to apply/reformulate/reinterpret them in a GFT context.

Condensed matter theory also provides specific examples of systems in which the collective behavior of the microscopic constituents in some hydrodynamic approximation gives rise to effective emergent geometries [15]. The collective variables of the fluid, in these background configurations, can be recast as the component functions of an *effective metric*. Moreover, the effective dynamics of perturbations (quasi-particles, themselves collective excitations of the fundamental constituents of the fluid) around the same background configurations take the form of matter field theories in curved spacetimes, with the effective metrics obtained from the collective background parameters of the fluid. Inspired by these results, we ask: assuming that a given GFT model describes the microscopic dynamics of a *discrete quantum space*, and that some solution of the corresponding fundamental equations can be interpreted as identifying a given quantum spacetime configuration (this is justified also by the interpretation of the GFT equations of motion, as we have discussed), can we obtain an effective macroscopic *continuum* field theory for matter fields from it, using a similar strategy? And if so, what is the effective spacetime and geometry that these emergent matter fields see?

The answer is affirmative [38, 54], and the effective matter field theories that we obtain most easily from GFTs are quantum field theories on non-commutative spaces of Lie algebra type.

The results in this direction are still quite scarce and still rather tentative, but they represent a starting point in a promising, in our opinion, research direction.

The basic point is the use of the natural duality between Lie algebra and corresponding Lie group, the same duality at the basis of the Lie algebra representation of GFTs, reinterpreted as the non-commutative version of the usual duality between coordinate and momentum space. If we have a non-commutative spacetime of Lie algebra type $[X_\mu, X_\nu] = C^\lambda_{\mu\nu} X_\lambda$, the corresponding momentum space is naturally identified with the corresponding Lie group, in such a way that the non-commutative coordinates X_μ act on it as (Lie) derivatives. From this perspective, we understand the origin of the spacetime non-commutativity to be the curvature of the corresponding momentum space, a sort of Planck-scale "co-gravity" [69]. The link with GFTs is then obvious: in momentum space the field theory on such non-commutative spacetime will be given, by definition, by some sort of group field theory. The task is to derive interesting matter field theories from interesting GFT models of quantum spacetime.

3D case

In three spacetime dimensions the results obtained concern a Euclidean non-commutative spacetime given by the $\mathfrak{su}(2)$ Lie algebra, i.e. whose spacetime coordinates are identified with the $\mathfrak{su}(2)$ generators with $[X_i, X_j] = i\frac{1}{\kappa}\varepsilon_{ijk}X_k$. Momenta are instead identified with group elements $\mathrm{SU}(2)$ [69], acquiring a non-commutative addition property following the group composition law. The duality is realized by the same non-commutative Fourier transform we introduced earlier. A scalar field theory in momentum space is then given by a group field theory of the type:

$$S[\psi] = \frac{1}{2}\int_{\mathrm{SU}(2)} dg\psi(g)\mathcal{K}(g)\psi(g^{-1}) - \frac{\lambda}{3!}\int [dg]^3\, \psi(g_1)\psi(g_2)\psi(g_3)\delta(g_1g_2g_3), \tag{12.24}$$

in the 3-valent case, where the integration measure is the Haar measure on the group, and $\mathcal{K}$ is some local kinetic term.

We also notice that the Feynman amplitudes of the above scalar field action (with simple kinetic terms) can be derived from the Ponzano–Regge spin foam model coupled to point particles [43], in turn obtainable from an extended GFT formalism [79]. We will see that the GFT construction to be presented allows us to bypass completely the spin foam formulation of the coupled theory.

Take now the Boulatov model we have introduced earlier, for a real field ϕ : $\mathrm{SU}(2)^3 \to \mathbb{R}$ invariant under the diagonal right action (we use right instead of left action to keep closer to the results as presented in [38]).

Now we look at 2-dimensional variations of the ϕ-field around classical solutions of the corresponding equations of motion:

$$\phi(g_3, g_2, g_1) = \frac{\lambda}{3!}\int dg_4dg_5dg_6\phi(g_3, g_4, g_5)\phi(g_5, g_2, g_6)\phi(g_6, g_4, g_1). \tag{12.25}$$

Calling $\phi^{(0)}$ a generic solution to this equation, we look at field perturbations $\delta\phi(g_1, g_2, g_3) \equiv \psi(g_1g_3^{-1})$ which do not depend on the group element g_2. We consider a specific class of classical solutions, named "flat" solutions (they can be interpreted as quantum flat space on some *a priori* non-trivial topology):

$$\phi^{(0)}(g_1, g_2, g_3) = \sqrt{\frac{3!}{\lambda}}\int dg\; \delta(g_1g)F(g_2g)\delta(g_3g), \quad F : G \to \mathbb{R}. \tag{12.26}$$

As shown in [38], this ansatz gives solutions to the field equations as soon as $\int F^2 = 1$.

This leads to an effective action for the 2D variations ψ:

$$\begin{aligned} S_{\mathrm{eff}}[\psi] = {} & \frac{1}{2}\int \psi(g)\mathcal{K}(g)\psi(g^{-1}) - \frac{\mu}{3!}\int [dg]^3\, \psi(g_1)\psi(g_2)\psi(g_3)\delta(g_1g_2g_3) \\ & -\frac{\lambda}{4!}\int [dg]^4\, \psi(g_1)..\psi(g_4)\delta(g_1..g_4), \end{aligned} \tag{12.27}$$

with the kinetic term and the 3-valent coupling given in terms of F:

$$\mathcal{K}(g) = 1 - 2\left(\int F\right)^2 - \int dh F(h)F(hg), \qquad \frac{\mu}{3!} = \sqrt{\frac{\lambda}{3!}}\int F,$$

with $F(g)$ assumed to be invariant under conjugation $F(g) = F(hgh^{-1})$.

Such an action defines a non-commutative quantum field theory invariant under the quantum double of SU(2) (a quantum deformation of the Poincaré group) [38, 43, 62, 66, 69].

Being an invariant function, F can be expanded in group characters:

$$F(g) = \sum_{j \in \mathbb{N}/2} F_j \chi_j(g), \qquad F_0 = \int F, \quad F_j = \int dg\, F(g)\chi_j(g), \qquad (12.28)$$

where the F_j's are the Fourier coefficients of the Peter–Weyl decomposition in irreducible representations of SU(2), labeled by $j \in \mathbb{N}/2$. The kinetic term then reads:

$$\mathcal{K}(g) = 1 - 3F_0^2 - \sum_{j \geq 0} \frac{F_j^2}{d_j} \chi_j(g) = \sum_{j \geq 0} F_j^2 \left(1 - \frac{\chi_j(g)}{d_j}\right) - 2F_0^2 \equiv Q^2(g) - M^2. \qquad (12.29)$$

It is easy to check that $Q^2(g) \geq 0$ with $Q(\mathbb{I}) = 0$. We interpret this term as the generalized "Laplacian" of the theory, while the 0-mode F_0 defines the mass $M^2 \equiv 2F_0^2$.

If we choose the simple solution (other choices will give more complicated kinetic terms)

$$F(g) = a + b\chi_1(g), \qquad \int F = a^2 + b^2 = 1, \qquad (12.30)$$

we obtain

$$\mathcal{K}(g) = \frac{4}{3}(1 - a^2)\, \vec{p}^{\,2} - 2a^2. \qquad (12.31)$$

4D case

The 4-dimensional non-commutative spacetime that is of most direct relevance for quantum gravity phenomenology is the so-called κ-Minkowski [4, 69]. We recall here some of its features, and refer to [54] for further details and references. κ-Minkowski spacetime can be identified with the Lie algebra $\mathfrak{an}_3$, which is a subalgebra of $\mathfrak{so}(4,1)$. Indeed, if $J_{\mu\nu}$ are the generators of $\mathfrak{so}(4,1)$, the $\mathfrak{an}_3$ generators are:

$$X_0 = \frac{1}{\kappa} J_{40}, \quad X_k = \frac{1}{\kappa}(J_{4k} + J_{0k}), \quad k = 1, \ldots, 3, \qquad (12.32)$$

which satisfy the commutation relations:

$$[X_0, X_k] = -\frac{i}{\kappa} X_k, \qquad [X_k, X_l] = 0, \quad k, l = 1, \ldots, 3. \qquad (12.33)$$

Using this, we can then define non-commutative plane waves with the AN_3 group elements as $h(k_\mu) = h(k_0, k_i) \equiv e^{ik_0 X_0} e^{ik_i X_i}$, thus identifying the coordinates on the group k_μ as the wave-vector (in turn related to the momentum). From here, a non-commutative addition of wave-vectors follows from the group multiplication of the corresponding plane waves.

Crucial for our construction is the Iwasawa decomposition [63], relating SO(4, 1) and AN_3 as:

$$\mathrm{SO}(4,1) = \mathrm{AN}_3\,\mathrm{SO}(3,1) \cup \mathrm{AN}_3\mathcal{M}\,\mathrm{SO}(3,1), \tag{12.34}$$

where the two sets are disjoint and $\mathcal{M}$ is the diagonal matrix with entries $(-1, 1, 1, 1, -1)$ in the fundamental 5D representation of SO(4, 1). Since de Sitter spacetime dS_4 can be defined as the coset $\mathrm{SO}(4,1)/\mathrm{SO}(3,1)$, an arbitrary point v on it can be obtained uniquely as:

$$v = (-)^\varepsilon h(k_\mu).v^{(0)} = h(k_\mu)\mathcal{M}^\varepsilon.v^{(0)}, \qquad \varepsilon = 0 \text{ or } 1, \quad h \in \mathrm{AN}_3, \tag{12.35}$$

where we have taken a reference spacelike vector $v^{(0)} \equiv (0,0,0,1) \in \mathbb{R}^4$, such that its little group is the Lorentz group SO(3, 1) and the action of SO(4, 1) on it sweeps the whole de Sitter space, and defined the vector $v \equiv h(k_\mu).v^{(0)}$ with coordinates:

$$v_0 = -\sinh\frac{k_0}{\kappa} + \frac{\mathbf{k}^2}{2\kappa^2}e^{k_0/\kappa}, \quad v_i = -\frac{k_i}{\kappa}, \quad v_4 = \cosh\frac{k_0}{\kappa} - \frac{\mathbf{k}^2}{2\kappa^2}e^{k_0/\kappa}.$$

The sign $(-)^\varepsilon$ corresponds to the two components of the Iwasawa decomposition. We then introduce the set $\mathrm{AN}_3^c \equiv \mathrm{AN}_3 \cup \mathrm{AN}_3\mathcal{M}$, such that the Iwasawa decomposition reads $\mathrm{SO}(4,1) = \mathrm{AN}_3^c\,\mathrm{SO}(3,1)$ and AN_3^c is isomorphic to the full de Sitter space. Actually, one can check that AN_3^c is itself a group. A crucial point is that the component v_4 of the above vector is left invariant by the action of the Lorentz group SO(3, 1). This suggests using this function of the "momentum" k_μ as a new (deformed) invariant energy–momentum (dispersion) relation, in the construction of a deformed version of particle dynamics and field theory on κ-Minkowski spacetime. This is the basis for much current QG phenomenology [4]. Finally, we will need an integration measure on AN_3 in order to define a Fourier transform. The choice of such measure is explained and motivated in [54]. For the free real scalar field $\phi : G \to \mathbb{R}$, we define the action

$$S(\phi) = \int dh\,\phi(h)\,\mathcal{K}(h)\,\phi(h), \quad \forall h \in G, \tag{12.36}$$

where dh is a left-invariant measure. We then interpret $G = \mathrm{AN}_3^c$ as the momentum space. We demand $\mathcal{K}(h)$ to be a function on G, invariant under the Lorentz transformations, which suggests using some function $\mathcal{K}(h) = f(v_4(h))$. Two common choices are

$$\mathcal{K}_1(h) = (\kappa^2 - \pi_4(h)) - m^2, \quad \mathcal{K}_2(h) = \kappa^2 - (\pi_4(h))^2 - m^2, \quad \pi_4 = \kappa v_4. \tag{12.37}$$

The above action is then Lorentz invariant if we choose a Lorentz-invariant measure dh_L.

Finally, the following generalized Fourier transform relates functions on the group $\mathcal{C}(G)$ and elements of the enveloping algebra $\mathcal{U}(\mathfrak{an}_3)$, i.e. non-commutative fields on the non-commutative spacetime $\mathfrak{an}_3$, i.e. on κ-Minkowski. For $G = \mathrm{AN}_3^c$,

$$\hat{\phi}(X) = \int_{\mathrm{AN}_3} dh_L^+ h(k_\mu)\,\phi^+(k) + \int_{\mathrm{AN}_3\mathcal{M}} dh_L^- h(k_\mu)\,\phi^-(k), \;\; X \in \mathfrak{an}_3,\; \hat{\phi}(X) \in \mathcal{U}(\mathfrak{an}_3),$$

where we used the non-abelian plane-wave $h(k_\mu)$ [41, 69]. The group field theory action on G can now be rewritten as a non-commutative field theory on κ-Minkowski (in the AN_3 case)

$$S(\phi) = \int dh_L\,\phi(h)\,\mathcal{K}(h)\,\phi(h) = \int d^4X\,\Big(\partial_\mu\hat{\phi}(X)\partial^\mu\hat{\phi}(X) + m^2\hat{\phi}^2(X)\Big). \tag{12.38}$$

The Poincaré symmetries are naturally deformed in order to be consistent with the non-trivial commutation relations of the κ-Minkowski coordinates [41]. The 4-dimensional case is less straightforward, and to some extent the results of the analysis are less satisfactory, at this stage. For a general 4D GFT related to topological BF quantum field theories with gauge group $\mathcal{G}$, as we have seen already, the action is:

$$\begin{aligned} S_{4d} = &\frac{1}{2}\int [dg]^4\,\phi(g_1,g_2,g_3,g_4)\phi(g_4,g_3,g_2,g_1) \\ &-\frac{\lambda}{5!}\int [dg]^{10}\phi(g_1,g_2,g_3,g_4)\phi(g_4,g_5,g_6,g_7)\phi(g_7,g_3,g_8,g_9) \\ &\times\phi(g_9,g_6,g_2,g_{10})\phi(g_{10},g_8,g_5,g_1), \end{aligned} \tag{12.39}$$

where the field is again required to be gauge-invariant, $\phi(g_1,g_2,g_3,g_4) = \phi(g_1g,g_2g,g_3g,g_4g)$ for any $g \in \mathcal{G}$.

We generalize to 4D the "flat solution" ansatz of the 3D group field theory as [38]:

$$\phi^{(0)}(g_i) \equiv \sqrt[3]{\frac{4!}{\lambda}}\int dg\,\delta(g_1g)F(g_2g)\widetilde{F}(g_3g)\delta(g_4g), \tag{12.40}$$

with $(\int F\widetilde{F})^3 = 1$.

A simpler special case of the classical solution above is obtained choosing $\widetilde{F}(g) = \delta(g)$ while keeping F arbitrary but with $F(\mathbb{I}) = 1$. Calling $c \equiv \int F$, the effective action becomes:

$$S_{\text{eff}}[\psi] = \frac{1}{2}\int \psi(g)\psi(g^{-1})\left[1-2c^2-2cF(g)F(g^{-1})\right]$$
$$-c\left(\sqrt[3]{\frac{\lambda}{4!}}\right)\int \psi(g_1)..\psi(g_3)\,\delta(g_1..g_3)\,[c+F(g_3)]$$
$$-c\left(\sqrt[3]{\frac{\lambda}{4!}}\right)^2\int \psi(g_1)..\psi(g_4)\,\delta(g_1..g_4) - \frac{\lambda}{5!}\int \psi(g_1)..\psi(g_5)\,\delta(g_1..g_5).$$

In order to make contact with deformed special relativity, we now specialize this construction to one that gives an effective field theory based on the momentum group manifold AN_3. We start then, from the group field theory describing SO(4, 1) BF theory.

From the quantum gravity perspective, there are several reasons of interest in this model. (1) The McDowell–Mansouri formulation (as well as related ones [50]) defines 4D gravity with a cosmological constant as a BF theory for SO(4, 1) plus a potential term which breaks the gauge symmetry from SO(4, 1) down to the Lorentz group SO(3, 1); this suggests trying to define quantum gravity in the spin foam context as a perturbation of a topological spin foam model for SO(4, 1) BF theory. These ideas could also be implemented directly at the GFT level, and the starting point would necessarily be a GFT for SO(4, 1) of the type we use here. (2) We expect [39, 76] any classical solution of this GFT model to represent quantum de Sitter space on some given topology, and such configurations would be physically relevant also in the non-topological case. (3) The spin foam/GFT model for SO(4, 1) BF theory seems the correct arena to build a spin foam model for 4D quantum gravity plus particles on de Sitter space [64], treating them as topological curvature defects for an SO(4, 1) connection, similarly to the 3D case [43].

Following the above procedure, we naturally obtain an effective field theory living on SO(4, 1). We want then to obtain from it an effective theory on $\mathrm{AN}_3^{\mathrm{c}}$. We choose:

$$F(g) = \alpha(v_4(g)+a)\vartheta(g), \qquad \widetilde{F}(g) = \delta(g). \tag{12.41}$$

The function v_4 is defined as a matrix element of g in the fundamental (non-unitary) 5-dimensional representation of SO(4, 1), $v_4(g) = \langle v^{(0)}|g|v^{(0)}\rangle$, where $v^{(0)} = (0,0,0,0,1)$ is, as previously, the vector invariant under the SO(3, 1) Lorentz subgroup. $\vartheta(g)$ is a cut-off function providing a regularization of F, so that it becomes an integrable function. Assuming that $\vartheta(\mathbb{I}) = 1$, we require $\alpha = (a+1)^{-1}$ in order for the normalization condition to be satisfied.

Then we can derive the effective action around such classical solutions for 2D field variations:

$$
\begin{aligned}
S_{\text{eff}}[\psi] = & \frac{1}{2}\int \psi(g)\psi(g^{-1})\left[1 - 2c^2 - \frac{2c\vartheta^2(g)(a+v_4(g))^2}{(a+1)^2}\right] \\
& -c\left(\frac{\lambda}{4!}\right)^{\frac{1}{3}} \int \psi(g_1)..\psi(g_3)\,\delta(g_1..g_3)\,[c+F(g_3)] \\
& -c\left(\frac{\lambda}{4!}\right)^{\frac{2}{3}} \int \psi(g_1)..\psi(g_4)\,\delta(g_1..g_4) - \frac{\lambda}{5!}\int \psi(g_1)..\psi(g_5)\,\delta(g_1..g_5),
\end{aligned}
$$

where $c = \int F$. Thus the last issue to address in order to properly define this action is to compute the integral of F; this can be done, and we refer to [54] for details.

We recognize the correct kinetic term for a DSR field theory. However, the effective matter field is still defined on a SO(4, 1) momentum manifold. The only remaining issue is therefore to understand the “localization" process of the field ψ to AN_3^c. This can be obtained in a variety of more or less satisfactory ways [54]. In any case, restricting to group elements $h_i \in \mathrm{AN}_3^c$, we have the field theory:

$$
\begin{aligned}
S_{\text{eff}}[\psi] = & \frac{1}{2}\int \psi(h)\psi(h^{-1})\left[1 - 2c^2 - 2cv_4(h)^2\vartheta(h)^2\right] \\
& -c\left(\frac{\lambda}{4!}\right)^{\frac{1}{3}} \int \psi(h_1)..\psi(h_3)\,\delta(h_1..h_3)\,[c+v_4(h_3)\vartheta(h_3)] \\
& -c\left(\frac{\lambda}{4!}\right)^{\frac{2}{3}} \int \psi(h_1)..\psi(h_4)\,\delta(h_1..h_4) - \frac{\lambda}{5!}\int \psi(h_1)..\psi(h_5)\,\delta(h_1..h_5),
\end{aligned}
$$

with implicit left-invariant measure on AN_3^c. We have thus derived a DSR scalar field theory with a κ-deformed Poincaré symmetry from the GFT for SO(4, 1) topological BF theory.

Notice that if we had started from a GFT model for BF theory with gauge group AN(3), we would have obtained the above result without any effort, like in the 3D case. However, we would have lost any direct connection with quantum gravity models, as it is unclear how such BF theory is related to gravity.

Work in this direction, including these preliminary results, is a step towards bridging the gap between our fundamental discrete theory of spacetime and the continuum description of spacetime we are accustomed to, thus bringing this class of models closer to quantum gravity phenomenology and experimental falsifiability. Moreover, contrary to the situation in analog gravity models in condensed matter, our GFT models are non-geometric and far from usual geometrodynamics in their formalism, but at the same time are expected to encode quantum geometric information and to determine, in particular in their classical solutions, a (quantum and therefore classical) geometry for spacetime. We are, in other words, far beyond a pure analogy, here.

12.5 Some important open issues

We list now, with no presumption of completeness, and in no particular order, several open issues in the GFT approach that we deem important (and some ideas for tackling them). We hope the brave and talented reader will pick them up and join our ongoing efforts to address them.

Obviously, the first issue is the *construction of a complete and convincing GFT model for 4D quantum gravity*.

From the point of view presented in this chapter, this means obtaining a GFT model whose Feynman amplitudes have a compelling expression as a simplicial path integral for 4D gravity, with a clear geometric meaning of the various contributions to the amplitude. Also, its spin foam expression should be derivable and possibly manageable, showing a clear translation of the geometric content of the amplitude in the algebraic language of group representation theory, at the quantum level. Finally, its space of boundary states should be clearly identified, as should its relation with the space of states of canonical loop quantum gravity.

This brings us to a second issue, which is to achieve *a rigorous and physically transparent link between the canonical LQG framework and the GFT/spin foam one*.

We would expect, in fact, that a proper covariant reformulation of the canonical dynamics of LQG quantum states, for any graph-changing Hamiltonian (or master) constraint operator [90] would lead to a GFT formalism, in the same way as the dynamics of any system of particles whose dynamics implies the possibility of creation and annihilation of the same is best captured in a QFT formalism. To realize this correspondence, one would need for example a rigorous definition and a detailed understanding of the space of GFT states as a Fock space, and to compare this with the kinematical state space of LQG, showing in what exact sense the first is a second quantization of the latter. Having this, one could attempt a derivation of the GFT path integral from LQG using coherent states, as often done in the usual QFT case. For this, the non-commutative GFT representation seems particularly suited. Any such derivation should lead to a better understanding of how any realistic LQG Hamiltonian/master constraint is encoded in GFT action, and thus in its equations of motion.

This brings us to the next issue, that of *solutions of the GFT equations of motion*.

Obviously, first of all it is necessary to identify more of them and understand their physical/geometric meaning (including the ones already found [38, 54]). Also, this should clarify in what sense they correspond to solutions of Hamiltonian/master constraints. A better control over the space of solutions of the classical GFT equations can be obtained, possibly, by learning to control and use the perturbative GFT tree-level expansion. Finally, this could be an avenue to investigate the physical (gravitational) meaning of GFT coupling constant.

The problem of gaining control over the tree-level expansion of the GFT is of course part of the more general issue, that we discussed above, of *gaining control and understanding of the GFT Feynman expansion*.

As we have shown, this is a complex matter, even for simple GFT models. It includes identifying the contribution of non-manifold configurations to the sum, and possibly suppressing it in some way. It involves addressing in full the problem of divergences: of individual Feynman amplitudes (simplicial path integrals/spin foam models), i.e. the perturbative GFT renormalization, and of the total perturbative sum, i.e. its (Borel) summability and non-perturbative definition. It means understanding the role of topology change and of its physical consequences, and controlling somehow the sum over topologies. Once more, work on these aspects of the GFT perturbative expansion could also shed light on the meaning of the GFT coupling constant, which has been suggested to govern the topology-changing processes [39, 76].

As we have mentioned, it is only recently [10] that some symmetries of discrete gravity, already identified at the level of the GFT amplitudes, i.e. at the level of simplicial gravity path integrals, have been identified also at the level of the corresponding GFT action. These are the local rotation and the translation symmetry that characterize BF theory, in any dimension. In particular, translation symmetry is crucial for the topological invariance of the theory, and it is also closely related to diffeomorphism symmetry, to the point that it could almost be identified with it [45]. Therefore, this result is a starting clue for a more extensive *analysis of the GFT symmetries and their relation with the symmetries of simplicial gravity; in particular, the identification of diffeomorphism in GFTs*. In particular, having identified diffeomorphism symmetry in 3D gravity, and translation symmetry in 4D BF, we have now to understand how exactly they are broken [31] when passing to 4D gravity. In fact, it can already be seen that the imposition of the gravity constraints on 4D BF theory breaks the BF translation symmetry. It remains however to study if there is any remnant of such symmetry, the details of its breaking, and how it can be recovered in some (still geometric) regime. The possibility of working in a simplicial path integral representation for the GFT amplitudes is going to be crucial. Having identified the symmetries of various GFT models, one should develop a systematic analysis of these symmetries (and others) and of their consequences at GFT level, using QFT tools (e.g. Ward identities). This analysis should also provide clues for the study of the continuum limit of the theory, which, if indeed it gives back some (modified) version of general relativity, is almost fully characterized by the presence of diffeomorphism symmetry.

And the problem of *the continuum approximation and the link with general relativity (and matter field theories)* is the real outstanding open issue, like in most other approaches to non-perturbative and background-independent quantum gravity.

Here there are two possibilities. The GFT perturbative expansion around the vacuum is an expansion in simplicial complexes of higher and higher complexity. The first possibility is that we can give a physical, continuum meaning to working with a simplicial complex, in terms of some precise truncation of the degrees of freedom of the full continuum theory, or in terms of some large-scale approximation. If this is the case, then it could be enough, at least for a subset of physically interesting questions, to work at a given low order in the GFT perturbation theory, and to compute physical quantities using a fixed simplicial complex, or a finite number of them. The questions is then: are a few simplices or some simple spin networks enough, in order to compute approximate, but physically meaningful *continuum* geometric quantities? If the answer is positive, then we can use many results in simplicial gravity to do so, as well as the coherent state techniques for fixed spin network graphs in LQG and spin foams, to try to extract physics from GFT models. A different possibility is that small numbers of simplices or simple graphs just cannot capture in any adequate way continuum information, especially concerning the dynamics of the theory. In order to do so, then, one would have to use highly refined simplicial complexes made out of very high numbers of simplices. From the GFT point of view, this means that the physics of continuum spacetimes has to be looked for in the regime of many GFT particles, thus far from the non-perturbative vacuum around which the GFT Feynman expansion is defined.

If this is the case, then one should develop a *statistical GFT, i.e. the statistical mechanics (or field theory) of many GFT quanta*.

In particular, the question of the continuum approximation becomes the question of identifying the correct *GFT phase(s), the corresponding phase transition(s), and the relevant regime(s) of dynamical variables* around which a continuum approximation becomes valid and the effective dynamics of the GFT system is described by (maybe modified) general relativity. Again, if the physics of the continuum spacetime is the physics of large ensembles of GFT quanta, then the correct conceptual strategy is to treat quantum space as a sort of weird condensed matter system with microscopic, atomic description given by some (class of) GFT models. Notice that this is exactly what happens in matrix models. The issue is then first of all to develop the appropriate mathematical tools to study the thermodynamic limit of GFTs, identify the relevant phases, and extract the effective GFT dynamics around them. Second, one needs to devise methods (and probably an appropriate conceptual framework) to re-express this effective dynamics in spacetime and geometric terms, i.e. from the GFT "pre-geometric" language to the language of continuum general relativity. We are already in a rather speculative setting, here. But we could speculate further [75] that GFTs could provide the right framework to realize the idea [60, 92] of continuum spacetime as a sort of condensate, in the precise form of a condensed or fluid phase of (very many) GFT quanta, simplices or spin network vertices, and of GR as a sort of hydrodynamics for these fundamental quanta of space in such a regime.

Whether these speculations are correct or not, and whether we can realize them fully and rigorously or not, the really important point is to be able to *extract physical predictions from GFTs, possibly together with a better understanding of the fundamental nature of space and time*, and say something new and interesting about our world. We have mentioned above one possible avenue towards this goal, i.e. deriving effective non-commutative field theories for matter fields from GFT. There are certainly many others. One could be making contact with loop quantum cosmology [22], in some sector of the GFT. More generally, any of the procedures we will identify for making contact with continuum physics, by statistical GFT or condensed matter techniques applied to GFT, and with GR, will also provide, after this link has been established, effective models with implications for phenomenology.

12.6 Conclusions

To conclude, we have introduced the key ideas behind the group field theory approach to quantum gravity, and to the tentative microscopics of quantum space it suggests. We have also introduced the basic elements of the GFT formalism, focusing on the 3-dimensional case for simplicity. We have reported briefly on the current status of the work devoted to the construction of interesting and viable 4-dimensional GFT models. Finally, we have also briefly reported on some recent results obtained in this approach, concerning both the mathematical definition of these models as bona fide field theories, and possible avenues towards extracting interesting physics from them. We hope that, from our outline, it shows clearly that, while much more work is certainly needed in this area of research, the new direction towards quantum gravity that group field theories provide is exciting and full of potential.

Acknowledgments

We thank A. Baratin, J. Ben Geloun, V. Bonzom, G. Calcagni, B. Dittrich, L. Freidel, F. Girelli, R. Gurau, E. Livine, R. Pereira, V. Rivasseau, C. Rovelli, J. Ryan, L. Sindoni, M. Smerlak, J. Tambornino, and many other colleagues for useful discussions and collaboration on the topics dealt with in this chapter. We thank the organizers of the EllisFest, and in particular George Ellis and Jeff Murugan, and the participants, for a very useful and enjoyable workshop. We also thank the editors of this volume for their patience with a difficult contributor. Support from the A. von Humboldt Stiftung through a Sofja Kovalevskaja Prize is gratefully acknowledged.

References

[1] S. Alexandrov, *Phys. Rev. D* **78**, 044033 (2008) [arXiv: 0802.3389 [gr-qc]].
[2] J. Ambjørn, B. Durhuus, T. Jonsson, *Mod. Phys. Lett.* **A6**, 1133–46 (1991).
[3] J. Ambjørn, J. Jurkiewicz, R. Loll, *Phys. Rev. D* **72**, 064014 (2005) [arXiv: hep-th/0505154].
[4] G. Amelino-Camelia, *Lect. Notes Phys.* **669**, 59–100 (2004) [arXiv: gr-qc/0412136].
[5] P. Anspinwall, B. Greene, D. Morrison, *Nucl. Phys. B* **416**, 414–80 (1994) hep-th/9309097.
[6] J. C. Baez, J. W. Barrett, *Adv. Theor. Math. Phys.* **3**, 815 (1999) gr-qc/9903060.
[7] T. Banks, *Nucl. Phys.* B **309**, 493 (1988).
[8] A. Baratin, C. Flori, T. Thiemann [arXiv: 0812.4055 [gr-qc]].
[9] A. Baratin, B. Dittrich, D. Oriti, J. Tambornino (2010), [arXiv:1004.3450 [hep-th]].
[10] A. Baratin, F. Girelli, D. Oriti (2010), [arXiv:1101.0590 [hep-th]].
[11] A. Baratin, D. Oriti [arXiv: 1002.4723 [hep-th]].
[12] A. Baratin, D. Oriti (2010), to appear.
[13] A. Baratin, D. Oriti (2010), *Phys. Rev. Lett.* **105** 221302 (2010).
[14] A. Barbieri, *Nucl. Phys. B* **518**, 714 (1998) gr-qc/9707010.
[15] C. Barcelo, S. Liberati, M. Visser, *Living Rev. Rel.* **8**, 12 (2005) [arXiv: gr-qc/0505065].
[16] J. W. Barrett, L. Crane, *J. Math. Phys.* **39**, 3296 (1998), gr-qc/9709028.
[17] J. Barrett, R. Dowdall, W. Fairbairn, H. Gomes, F. Hellman, *J. Math. Phys.* **50**, 112504 (2009), [arXiv:0902.1170 [gr-qc]].
[18] J. Barrett, R. Dowdall, W. Fairbairn, F. Hellman, R. Pereira, *Class. Quant. Grav.* **27**, 165009 (2010) [arXiv: 0907.2440 [gr-qc]].
[19] J. Barrett, I. Naish-Guzman, *Class. Quant. Grav.* **26**, 155014 (2009) [arXiv: 0803.3319 [gr-qc]].
[20] J. Ben Geloun, J. Magnen, V. Rivasseau *Euro. Phys. J.* **C70**, 1119–30 (2010) [arXiv: 0911.1719 [hep-th]].
[21] J. Ben Geloun, T. Krajewski, J. Magnen, V. Rivasseau *Class. Quant. Grav.* **27**, 155012 (2010) [arXiv: 1002.3592 [hep-th]].
[22] M. Bojowald, *Living Rev. Rel.* **11**, 4 (2008).

[23] V. Bonzom, E. Livine *Phys. Rev.* **D79**, 064034 (2009) [arXiv:0812.3456].
[24] D. V Boulatov, *Mod. Phys. Lett.* **A7**, 1629–46 (1992) [arXiv:hep-th/9202074].
[25] S. Coleman, *Nucl. Phys. B* **310**, 643 (1988).
[26] F. Conrady, L. Freidel, *Phys. Rev. D* **78**, 104023 (2008) [arXiv: 0809.2280].
[27] F. Conrady, L. Freidel, *Class. Quant. Grav.* **25**, 245010 (2008) [arXiv: 0806.4640].
[28] F. David, *Nucl. Phys. B* **257**, 45 (1985).
[29] R. De Pietri, L. Freidel, *Class. Quant. Grav.* **16**, 2187 (1999) gr-qc/9804071.
[30] R. De Pietri, L. Freidel, K. Krasnov, C. Rovelli, *Nucl. Phys. B* **574**, 785 (2000) [arXiv: hep-th/9907154].
[31] B. Dittrich, [arXiv:0810.3594[gr-qc]].
[32] B. Dittrich, J. Ryan *Phys. Rev.* **D82**, 064026 (2010) [arXiv:0807.2806].
[33] F. Dowker, R. Sorkin, *Class. Quant. Grav.* **15**, 1153–67 (1998) gr-qc/9609064.
[34] F. Dowker, in *The Future of Theoretical Physics and Cosmology*, 436–52, Cambridge University Press (2002), gr-qc/0206020.
[35] J. Engle, R. Pereira, C. Rovelli, *Phys. Rev. Lett.* **99**, 161301 (2007) [arXiv: 0705.2388].
[36] J. Engle, R. Pereira, C. Rovelli, *Nucl. Phys. B* **798**, 251 (2008) [arXiv: 0708.1236].
[37] J. Engle, E. Livine, R. Pereira, C. Rovelli, *Nucl. Phys. B* **799**, 136 (2008) [arXiv:0711.0146].
[38] W. Fairbairn, E. Livine, *Class. Quant. Grav.* **24**, 5277 (2007) [arXiv: gr-qc/0702125].
[39] L. Freidel, *Int. J. Phys.* **44**, 1769–83, (2005) [arXiv: hep-th/0505016].
[40] L. Freidel, R. Gurau, D. Oriti, *Phys. Rev. D* **80**, 044007 (2009) [arXiv: 0905.3772].
[41] L. Freidel, J. Kowalski-Glikman, S. Nowak (2007), arXiv:0706.3658 [hep-th].
[42] L. Freidel, K. Krasnov, *Class. Quant. Grav.* **25**, 125018 (2008) [arXiv: 0708.1595].
[43] L. Freidel, E. Livine, *Class. Quant. Grav.* **23**, 2021 (2006) [arXiv: hep-th/0502106].
[44] L. Freidel, E. Livine, C. Rovelli, *Class. Quant. Grav.* **20**, 1463–78 (2003) [arXiv: gr-qc/0212077].
[45] L. Freidel, D. Louapre, *Phys. Rev. D* **68**, 104004 (2003) [arXiv: hep-th/0211026].
[46] L. Freidel, D. Louapre, *Nucl. Phys. B* **662**, 279–98, 2003 [arXiv: gr-qc/0212001].
[47] L. Freidel, D. Louapre, *Class. Quant. Grav.* **21**, 5685–726 (2004) [arXiv: hep-th/0401076].
[48] L. Freidel, D. Louapre [arXiv: gr-qc/0410141].
[49] L. Freidel, S. Majid, Class. *Quant. Grav.* **25**, 045006 (2008) [arXiv:hep-th/0601004].
[50] L. Freidel, A. Starodubtsev (2005) arXiv:hep-th/0501191.
[51] M. Gaul, C. Rovelli, *Lect. Notes Phys*. **541**, 277 (2000) gr-qc/9910079.
[52] S. Giddings, A. Strominger, *Nucl. Phys. B* **321**, 481 (1989).
[53] P. Ginsparg, “Matrix models of 2-d gravity”, [arXiv: hep-th/9112013].
[54] F. Girelli, E. Livine, D. Oriti, *Phys. Rev. D* **81**, 024015 (2010) [arXiv: 0903.3475 [gr-qc]].
[55] D. Giulini, *Gen. Rel. Grav.* **41**, 785–815 (2009) [arXiv:0902.3923].
[56] M. Gross, *Nucl. Phys. Proc. Suppl.* **25A**, 144–149 (1992).
[57] R. Gurau [arXiv:0907.2582 [hep-th]].
[58] R. Gurau [arXiv:0911.1945 [hep-th]].
[59] G. Horowitz, *Class. Quant. Grav.* **8**, 587–602 (1991).
[60] B. L. Hu, *Int. J. Theor. Phys.* **44** (2005) 1785–806 [arXiv:gr-qc/0503067].
[61] C. Isham, gr-qc/9510063.
[62] E. Joung, J. Mourad, K. Noui, *J. Math. Phys.* **50**, 052503 (2009) [arXiv:0806.4121 [hep-th]].
[63] A. Klimyk, N. Vilenkin, *Representations of Lie Groups and Special Functions*, Springer Ed. (1995).
[64] J. Kowalski-Glikman, A. Starodubtsev, *Phys. Rev. D* **78**, 084039 (2008), arXiv:0808.2613.

[65] K. Kuchar, in Winnipeg 1991, Proceedings, General Relativity and Relativistic Astrophysics, pp. 211–314.
[66] E. Livine *Class. Quant. Grav.* **26**, 195014 (2009) [arXiv:0811.1462 [gr-qc]].
[67] E. Livine, S. Speziale, *Europhys. Lett.* **81**, 50004 (2008) [arXiv:0708.1915 [gr-qc]].
[68] J. Magnen, K. Noui, V. Rivasseau, M. Smerlak, *Class. Quant. Grav.* **26**, 185012 (2009) [arXiv:0906.5477].
[69] S. Majid, *Foundations of Quantum Group Theory*, Cambridge University Press (1995).
[70] M. McGuigan, *Phys. Rev. D* **38**, 3031 (1988).
[71] H. Ooguri, *Mod. Phys. Lett.* **A7**, 2799 (1992) hep-th/9205090.
[72] D. Oriti, *Rept. Prog. Phys.* **64**, 1489 (2001) [arXiv: gr-qc/0106091].
[73] D. Oriti, *Phys. Lett. B* **532**, 363–72 (2002) [arXiv: gr-qc/0201077].
[74] D. Oriti, PhD thesis, University of Cambridge (2003) [arXiv: gr-qc/0311066].
[75] D. Oriti, in *Quantum Gravity*, B. Fauser, J. Tolksdorf, E. Zeidler (eds.), Birkhaeuser, Basel (2007) [arXiv: gr-qc/0512103].
[76] D. Oriti [arXiv:gr-qc/0607032].
[77] D. Oriti, Proceedings of Science [arXiv:0710.3276].
[78] D. Oriti (ed.), *Approaches to Quantum Gravity*, Cambridge University Press, Cambridge (2009).
[79] D. Oriti, J. Ryan, *Class. Quant. Grav.* **23**, 6543 (2006) [arXiv: gr-qc/0602010].
[80] A. Perelomov, *Generalized Coherent States and their Applications*, Springer, Berlin (1986).
[81] A. Perez, *Class. Quant. Grav.* **20**, R43 (2003) [arXiv: gr-qc/0301113].
[82] A. Perez, C. Rovelli, *Nucl. Phys. B* **599**, 255 (2001) [arXiv: gr-qc/0006107].
[83] M. P. Reisenberger, gr-qc/9804061.
[84] C. Rovelli, in the Proceedings of the 9th Marcel Grossmann Meeting, Rome, Italy (2000), V. G. Gurzadyan *et al*. (eds), Singapore, World Scientific, gr-qc/0006061.
[85] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge (2006).
[86] L. Smolin, in D. Rickles (ed.), *The Structural Foundations of Quantum Gravity*, pp. 196–239, hep-th/0507235.
[87] R. Sorkin, *Int. J. Theor. Phys.* **30**, 923–48 (1991).
[88] R. Sorkin, *Int. J. Theor. Phys.* **36**, 2759–81 (1997) gr-qc/9706002.
[89] C. Teitelboim, *Phys. Rev. D* **25**, 3159 (1982).
[90] T. Thiemann, *Modern Canonical Quantum general Relativity*, Cambridge University Press, Cambridge (2007).
[91] V. Turaev, O. Viro, *Topology* **31**, 865 (1992).
[92] G. E. Volovik, Proceedings of MG11, session “Analog Models of and for General Relativity”, arXiv:gr-qc/0612134.
[93] R. Williams, in [78].

13

Causal dynamical triangulations and the quest for quantum gravity

J. AMBJØRN, J. JURKIEWICZ & R. LOLL

Quantum gravity by causal dynamical triangulation has over the last few years emerged as a serious contender for a nonperturbative description of the theory. It is a nonperturbative implementation of the sum-over-histories, which relies on few ingredients and initial assumptions, has few free parameters and – crucially – is amenable to numerical simulations. It is the only approach to have demonstrated that a classical universe can be generated dynamically from Planckian quantum fluctuations. At the same time, it allows for the explicit evaluation of expectation values of invariants characterizing the highly nonclassical, short-distance behaviour of spacetime. As an added bonus, we have learned important lessons on which aspects of spacetime need to be fixed a priori as part of the background structure and which can be expected to emerge dynamically.

13.1 Quantum gravity – taking a conservative stance

Many fundamental questions about the nature of space, time and gravitational interactions are not answered by the classical theory of general relativity, but lie in the realm of the still-searched-for *theory of quantum gravity*: What is the quantum theory underlying general relativity, and what does it say about the quantum origins of space, time and our universe? What is the microstructure of spacetime at the shortest scale usually considered, the Planck scale $\ell_{\mathrm{Pl}} = 10^{-35}$ m, and what are the relevant degrees of freedom determining the dynamics there? Are they the geometric dynamical variables of the classical theory (or some short-scale version thereof), or do they also include the topology and/or dimensionality of

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

spacetime, quantities that classically are considered fixed? Can the dynamics of these microscopic degrees of freedom *explain* the observed large-scale structure of our own universe, which resembles a de Sitter universe at late times? Do notions like ‘space’, ‘time’ and ‘causality’ remain meaningful on short scales, or are they merely macroscopically *emergent* from more fundamental, underlying Planck-scale principles?

Despite considerable efforts over the last several decades, it has so far proven difficult to come up with a consistent and quantitative theory of quantum gravity, which would be able to address and answer such questions [31]. In the process, researchers in high-energy theory have been led to consider ever more radical possibilities in order to resolve this apparent impasse, from postulating the existence of extra structures unobservable at low energies to invoking ill-defined ensembles of multiverses and anthropic principles [23]. A grand unified picture has quantum gravity inextricably linked with the quantum dynamics of the three other known fundamental interactions, which requires a new unifying principle. Superstring theory is an example of such a framework, which needs the existence of an as yet unseen symmetry (supersymmetry) and ingredients (strings, branes, fundamental scalar fields). Loop quantum gravity, a non-unified approach, postulates the existence of certain fundamental quantum variables of Wilson loop type. Even more daring souls contemplate – inspired by quantum-gravitational problems – the abandonment of locality [25] or substituting quantum mechanics with a more fundamental, deterministic theory [28].

In view of the fact that none of these attempts has as yet thrown much light on the questions raised above, and that we have currently neither direct tests of quantum gravity nor experimental facts to guide our theory-building, a more conservative approach may be called for. What we will sketch in the following is an alternative route to quantum gravity, which relies on nothing but standard principles from quantum field theory, and on ingredients and symmetries already contained in general relativity. Its main premise is that *the framework of standard quantum field theory is sufficient to construct and understand quantum gravity as a fundamental theory,* if *the dynamical, causal and nonperturbative properties of spacetime are taken into account properly.*

Significant support for this thesis comes from a new candidate theory, *quantum gravity from causal dynamical triangulation (CDT)*, whose main ideas and results will be described below. CDT quantum gravity is a nonperturbative implementation of the gravitational path integral, and has already passed a number of nontrivial tests with regard to producing the correct classical limit. Its key underlying idea was conceived more than 10 years ago [1], in an effort to combine the insights of geometry-based nonperturbative canonical quantum gravity with the powerful calculational and numerical methods available in covariant approaches. After several years of modelling and testing both the idea and its implementation in spacetime dimensions two and three, where they give rise to nontrivial dynamical systems of 'quantum geometry' [5, 6], the first results for the physically relevant case of four dimensions were published in 2004 [7, 9].

Let us also mention that an independent approach to the quantization of gravity, much in the spirit of our main premise[1] and based on the 30-year-old idea of 'asymptotic safety' has been developing over roughly the same time period [39, 40]. It shares some features (covariance, amenability to numerical computation) as well as some results (on the spectral dimension) with CDT quantum gravity, and may ultimately turn out to be related.

13.2 What CDT quantum gravity is about

Quantum gravity theory based on causal dynamical triangulations is an explicit, nonperturbative and background-independent realization of the formal *gravitational path integral* (a.k.a. the 'sum over histories') on a differential manifold M,

$$Z(G_N, \Lambda) = \int_{\mathcal{G}(\mathrm{M}) = \frac{\mathrm{Lor(M)}}{\mathrm{Diff\,(M)}}} \mathcal{D}g_{\mu\nu}\; \mathrm{e}^{iS^{\mathrm{EH}}[g_{\mu\nu}]}, \quad S^{\mathrm{EH}} = \int d^4x \sqrt{\det g}\,(\frac{1}{G_N} R - 2\Lambda), \tag{13.1}$$

where S^{EH} denotes the four-dimensional Einstein–Hilbert action, G_N is the gravitational or Newton's constant, and Λ the cosmological constant. The path integral is to be taken over all spacetimes (metrics $g_{\mu\nu}$ modulo diffeomorphisms), with specified boundary conditions. The method of causal dynamical triangulation turns (13.1) into a well-defined finite and regularized expression, which can be evaluated and whose continuum limit (removal of the regulator) can be studied systematically [4].

One proceeds in analogy with the path-integral quantization à la Feynman and Hibbs of the nonrelativistic particle. This is defined as the continuum limit of a regularized sum over paths, where the contributing 'virtual' paths are taken from an ensemble of piecewise straight paths, with the length a of the individual segments going to zero in the limit. The corresponding CDT prescription in higher dimensions is to represent the space $\mathcal{G}$ of all Lorentzian spacetimes in terms of a set of triangulated, piecewise flat manifolds,[2] as originally introduced in the classical theory as 'general relativity without coordinates' [42]. For our purposes, the simplicial approximation $\mathcal{G}_{a,N}$ of $\mathcal{G}$ contains all simplicial manifolds T obtained

[1] Although the role of 'causality', which enters crucially in CDT quantum gravity, remains unclear in this approach.
[2] Unlike in the particle case, there is no embedding space; all geometric spacetime data are defined intrinsically, just like in the classical theory.

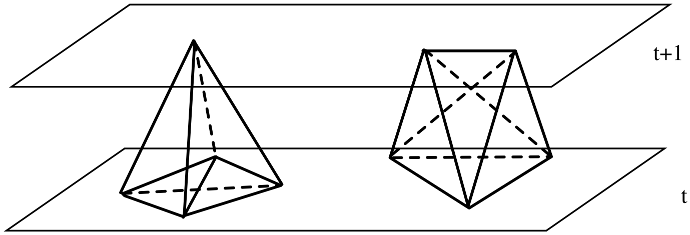


Figure 13.1 The two fundamental building blocks of CDT are four-simplices with flat, Minkowskian interior. They are spanned by spacelike edges, which lie entirely within spatial slices of constant time t, and timelike edges, which interpolate between adjacent slices of integer time. A building block of type (m,n) has m of its vertices in slice t, and n in slice $t+1$.

from gluing together at most N four-dimensional, triangular building blocks of typical edge length a, with a again playing the role of an ultraviolet (UV) cut-off (see Figure 13.1). The explicit form of the regularized gravitational path integral in CDT is

$$Z_{a,N}^{\mathrm{CDT}} = \sum_{\substack{\text{triangulated} \\ \text{spacetimes } T \in \mathcal{G}_{a,N}}} \frac{1}{C_T} \mathrm{e}^{i S^{\mathrm{Regge}}[T]}, \tag{13.2}$$

where S^{Regge} is the Regge version of the Einstein–Hilbert action associated with the simplicial spacetime T, and C_T denotes the order of its automorphism group. The discrete volume N acts as a volume cut-off. We still need to consider a suitable continuum or scaling limit

$$Z^{\mathrm{CDT}} := \lim_{\substack{N\to\infty \\ a\to 0}} Z_{a,N}^{\mathrm{CDT}} \tag{13.3}$$

of (13.2), while renormalizing the original bare coupling constants of the model, in order to arrive (if all goes well) at a theory of quantum gravity. The two limits in (13.3) are usually tied together by keeping a physical four-volume, defined as $V_4 := a^4 N$, fixed. In the limiting process a is taken to zero, $a \to 0$, and the individual discrete building blocks are then literally 'shrunk away'.

Let us summarize the key features of the construction scheme thus introduced. Unlike what is possible in the continuum theory, the path integral (13.2) is defined directly on the physical configuration space of *geometries*. It is nonperturbative in the sense of including geometries which are 'far away' from any classical solutions, and it is background-independent in the sense of performing the sum 'democratically', without distinguishing any given geometry (say, as a preferred background). Of course, these nice properties of the regularized path integral are only useful because *we are able to evaluate* Z^{CDT} *quantitatively*, with an essential role being played by Monte Carlo simulations. These, together with the associated finite-size scaling techniques [38], have enabled us to extract information about the nonperturbative, strongly coupled quantum dynamics of the system which is currently not accessible by analytical methods, neither in this nor any other approach to quantum gravity. It is reminiscent of the role played by lattice simulations in pinning down the nonperturbative behaviour of QCD (although this is a theory we already know *much* more about than quantum gravity).

13.3 What CDT quantum gravity is not about

Although causal dynamical triangulation is sometimes called a discrete approach, this is potentially misleading. First, one can of course think of the simplicial building blocks as discrete objects, but they are assembled into spacetimes that are perfectly continuous and not discrete. The space of geometries *is* discretized in the sense that both four-volume and curvature contribute in discrete 'bits' to the total action. However, this is only a feature of the chosen regularization, and has no physical significance as such. As explained in the previous section, the characteristic edge length a plays the role of an intermediate regulator and UV cut-off for the geometry. In the continuum limit, a is to be taken to zero strictly. In practice, what will usually suffice is to choose a significantly smaller than the scale at which one is trying to extract physical results, hence $a \ll \ell_{\mathrm{Pl}}$ if we want to establish Planck-scale dynamics.

Adherents of the idea of fundamental discreteness might be tempted to identify the edge length a with a fundamental, shortest length scale, typically, the Planck length. However, this would be an ad hoc prescription which is in no way required by the construction. Besides, it has the unpleasant feature that physics at the Planck scale will then depend explicitly on the details of the chosen regularization. For example, choosing squares instead of triangles, or choosing a different discrete realization of the Einstein action, will in general lead to different Planckian dynamics, thus introducing an infinite ambiguity *at that scale*. It is not good enough if all these different theories produce identical classical physics on large scales, because in quantum gravity one is of course interested in finding a (hopefully unique) description of physics at the Planck scale.

Instead of putting it in by hand, the issue of fundamental discreteness in quantum gravity needs to be addressed *dynamically*. Is such a scale generated by the dynamics of the theory? Although there are numerous claims that Planck-scale discreteness is almost 'self-evident' (often, to render one's favourite calculation of black hole entropy finite), there is at this stage no concrete evidence for such a discreteness in full, four-dimensional quantum gravity.[3] We have up to now not seen any indication of it, but it is conceivable that there exist nonperturbative quantum operators in CDT quantum gravity which measure lengths (or higher-dimensional volumes) and have a discrete spectrum as $a \to 0$, thus indicating fundamental discreteness. Even if such a discreteness were found, whether or not the currently unknown 'fundamental excitations of quantum gravity' are discrete or not may be yet another issue. It is not even particularly clear what one means by such a statement and whether it can be turned into an operationally well-defined question in the nonperturbative theory, and not one which is merely a feature of a particular representation of the quantum theory.

As we will see in more detail below, CDT is – as far as we are aware – the only nonperturbative approach to quantum gravity which has been able to dynamically generate its own, physically realistic background from nothing but quantum fluctuations. More than that, because of the minimalist set-up and the methodology used (quantum field theory and critical phenomena), the results obtained are robust in the sense of being largely independent of the details of the chosen regularization procedure and containing few free parameters. This is therefore also one of the perhaps rare instances of a candidate theory of quantum gravity which can potentially be falsified. In fact, the Euclidean version of the theory extensively studied in the 1990s has already been falsified because it does not lead to the correct classical limit [18, 22]. CDT quantum gravity improves on this previous attempt by building a causal structure right into the fabric of the model. Our investigations of both the quantum properties and the classical limit of this candidate theory are at this stage not sufficiently complete to provide conclusive evidence that we have found *the* correct theory of quantum gravity, but results until now have been unprecedented and most encouraging, and have thrown up a number of nonperturbative surprises.

13.4 CDT key achievements I – demonstrating the need for causality

We will confine ourselves to highlighting some of the most important results and new insights obtained in CDT quantum gravity, without entering into any of the technical details. The reader is referred to the literature cited in the text, as well as to the various overview articles available on the subject [15] for more information.

The crucial lesson learned for nonperturbative gravitational path integrals from CDT quantum gravity is that the ad hoc prescription of integrating over curved Euclidean *spaces* of metric signature (++++) instead of the physically correct curved Lorentzian space*times* of metric signature (−+++) generally leads to inequivalent

[3] The derivation of discrete aspects of the spectrum of the area and volume operators in loop quantum gravity [34, 43] disregards dynamics (in the form of the Hamiltonian constraint), quite apart from the fact that one can argue that discreteness has been put in 'by hand' by choosing a quantum representation where one-dimensional Wilson loops are well-defined operators.

and (in $d=4$) incorrect results. 'Euclidean quantum gravity' of this kind, as advocated by S. Hawking and collaborators [24], adopts this version of doing the path integral mainly for the technical reason to be able to use real weights $\exp(-S^{\mathrm{eu}})$ instead of the complex amplitudes $\exp(i\,S^{\mathrm{lor}})$ in its evaluation. The same prescription is used routinely in perturbative quantum field theory on flat Minkowski space, but in that case one can rely on the existence of a well-defined Wick rotation to relate correlation functions in either signature. This is *not* available in the context of continuum gravity beyond perturbation theory on a Minkowski background, but one may still hope that by *starting out* in Euclidean signature and quantizing this (wrong) theory, an inverse Wick rotation would then 'suggest itself' to translate back the final result into physical, Lorentzian signature. Alas, this has never happened, because – we would contend – no one has been able to make much sense of nonperturbative Euclidean quantum gravity in the first place, even in a reduced, cosmological context.[4]

CDT quantum gravity has provided the first explicit example of a nonperturbative gravitational path integral (in a toy model of two-dimensional gravity) which is exactly soluble and leads to distinct and inequivalent results depending on whether the sum over histories is taken over Euclidean spaces or Lorentzian spacetimes (or, more precisely, Euclidean spaces which are obtained by Wick rotation – which *does* exist for the class of simplicial spacetimes under consideration – from Lorentzian spacetimes). The Lorentzian path integral was first solved in [1], and a quantity one can compute and compare with the Euclidean version found in [2] is the cylinder amplitude (Figure 13.2). In the Lorentzian CDT case, only those histories are summed over which possess a global time slicing *with respect to which no spatial topology changes are allowed to occur*. After Wick rotation, this set constitutes a strict subset of all Euclidean (triangulated) spaces. In the latter there is no natural notion of 'time' or 'causality', and branching geometries are thus always present.

In the two-dimensional setting this means that one has identified a new class of anisotropic statistical mechanical models of fluctuating geometry. Several intriguing results that have been found from numerical simulations of matter-coupled versions of the model (which have so far resisted analytical solution) indicate that their geometric disorder is less severe than that of their Euclidean counterparts, and in particular that they seem to lead to critical matter exponents identical to those of the corresponding matter model on fixed, flat lattices [3, 12].

Another new direction in which the two-dimensional CDT model has been generalized is a controlled relaxation of the ban on branching points, while adhering to a global notion of proper time [35, 36]. This has culminated recently in the formulation of a fully-fledged CDT string field theory in zero target space dimensions [11].

[4] A discussion of the kind of problems that arise can be found in [26].

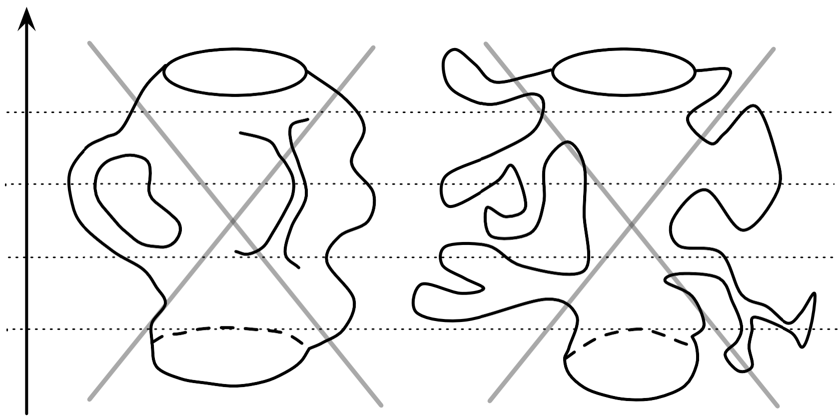
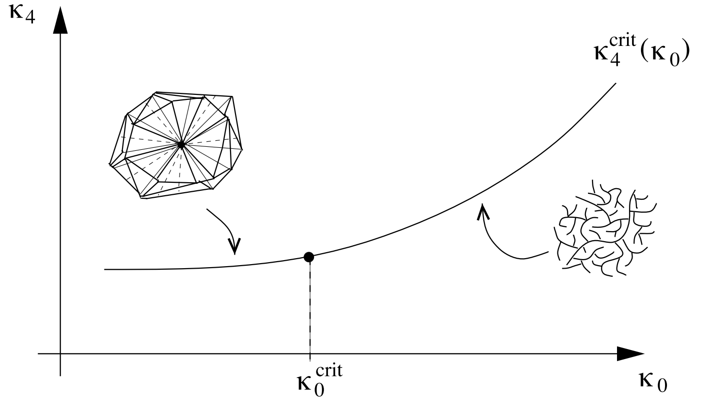


Figure 13.2 Typical history contributing to the loop–loop correlator in the 2D Lorentzian CDT path integral (left), time t is pointing up. The essential difference with the corresponding Euclidean amplitude is that the (one-dimensional) spatial slices, although quantum-fluctuating, are not allowed to change topology as a function of time t, thus avoiding causality-violating branching and merging points. This excludes spaces with wormholes (right of picture, a) and those with 'baby universes' branching out in the time direction (right of picture, b).

The matrix model formulation of the theory makes it possible to perform the sum over two-dimensional topologies explicitly [16]. These developments are described in more detail elsewhere in this volume [17].

Returning to the implementation of strict causality on path integral histories, a key finding of CDT quantum gravity is that a result similar to that found in two dimensions also holds in dimension four. The geometric degeneracy of the phases (in the sense of statistical systems) found in Euclidean dynamical triangulations [21, 22], and the resulting absence of a good classical limit, can in part be traced to the 'baby universes' present in the Euclidean approach also in four dimensions. As demonstrated by the results in [7, 9], the requirement of microcausality (absence of causality-violating points) of the individual path integral histories leads to a qualitatively new phase structure, containing a phase where the universe on large scales is extended and four-dimensional (Figure 13.3), as required by classical general relativity. Apart from the nice result that the problems of the Euclidean approach are cured by this prescription, this reveals an intriguing relation between the microstructure of spacetime (microcausality = suppression of baby universes in the time direction at sub-Planckian and bigger scales) and its emergent macrostructure. Referring to the questions raised at the beginning of Section 13.1, the more general lessons learned from this are that (i) 'causality' is not emergent, but needs to



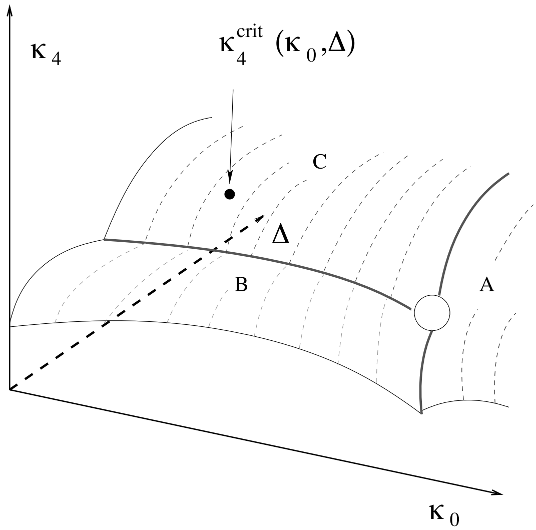


Figure 13.3 The phase diagrams of Euclidean (top) and Lorentzian (bottom) quantum gravity from dynamical triangulations, with κ_0 and κ_4 denoting the bare inverse Newton's constant and (up to an additive shift) the bare cosmological constant. After fine-tuning to the respective subspace where the cosmological constant is critical (tantamount to performing the infinite-volume limit), there are (i) two phases in EDT: the crumpled phase $\kappa_0 < \kappa_0^{\text{crit}}$ with infinite Hausdorff dimension and the branched-polymer phase $\kappa_0 > \kappa_0^{\text{crit}}$ with Hausdorff dimension 2, none of them with a good classical limit, (ii) three phases in CDT: A and B (the Lorentzian analogues of the branched-polymer and crumpled phases), and a *new* phase C, where an extended, four-dimensional universe emerges. The parameter Δ in CDT parametrizes a finite relative scaling between space- and time-like distances which is naturally present in the Lorentzian case.

be put in by hand on each spacetime history, and (ii) similarly, 'time' is not emergent. It is put into CDT by choosing a preferred (proper-)time slicing at the regularized level, but this turns out to be only a necessary condition to have a notion of time (as part of an extended universe) present in the continuum limit, at least on large scales. It is not sufficient, because in other phases of the CDT model (Figure 13.3) the spatial universe apparently does not persist at all (B) or only intermittently (A), see also [9].

13.5 CDT key achievements II – the emergence of spacetime as we know it

This brings us straight to the nature of the extended spacetime found in phase C of CDT quantum gravity. What is it, and how do we find out? We cannot just 'look at' the quantum superposition of geometries, which individually of course get wilder and spikier as the continuum limit $a \to 0$ is approached, just like the nowhere-differentiable paths of the path integral of the nonrelativistic particle [41]. We need to define and measure geometric *quantum observables*, evaluate their expectation values on the ensemble of geometries, and draw conclusions about the behaviour of the 'quantum geometry' generated by the computer simulations (that is, the ground state of minimal Euclidean action). Rather strikingly, inside phase C the many microscopic building blocks superposed in the nonperturbative path integral 'arrange themselves' into an extended quantum spacetime whose macroscopic shape is that of the well-known *de Sitter universe* [10,13]. This amounts to a highly nontrivial test of the classical limit, which is notoriously difficult to achieve in models of nonperturbative quantum gravity. The precise dynamical mechanism by which this happens is unknown, however, it is clear that 'entropy' (in other words, the measure of the path integral, or the number of times a given weight factor $\exp(-S)$ is realized) plays a crucial role in producing the outcome. This is reminiscent of phenomena in condensed matter physics, where systems of large numbers of microscopic, interacting constituents exhibit macroscopic, 'emergent' behaviour which is difficult to derive from the microscopic laws of motion. This makes it appropriate to think of CDT's de Sitter space as a *self-organizing quantum universe* [14].

The manner in which we have identified (Euclidean) de Sitter space from the computer data is by looking at the expectation value of the volume profile $V_3(t)$, that is, the size of the spatial three-volume as a function of proper time t. For a classical Lorentzian de Sitter space this is given by

$$V_3(t) = 2\pi^2 \left(c \cosh \frac{t}{c} \right)^3, \quad c = \text{const.} > 0, \tag{13.4}$$

which for $t > 0$ gives rise to the familiar, exponentially expanding universe, thought to give an accurate description of our own universe at late times, when matter can

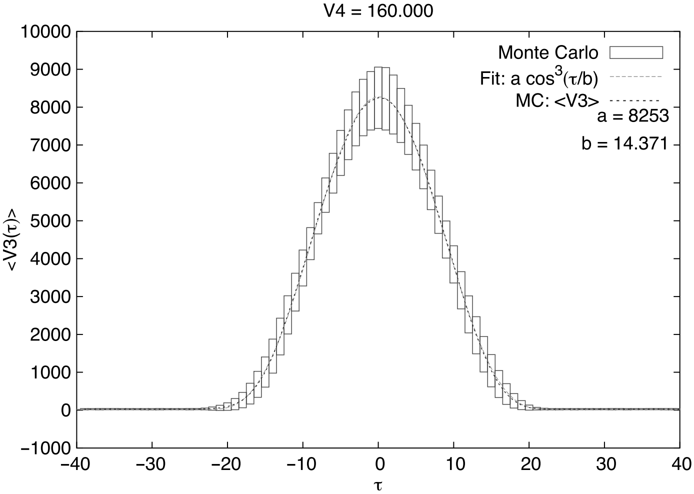


Figure 13.4 The shape $\langle V_3(\tau)\rangle$ of the CDT quantum universe, fitted to that of Euclidean de Sitter space (the 'round four-sphere') with rescaled proper time, $\langle V_3(\tau)\rangle = a\cos^3(\tau/b)$. Measurements taken for a universe of four-volume $V_4 =$ 160.000 and time extension $T = 80$. The fit of the Monte Carlo data to the theoretical curve for the given values of a and b is impressive. The vertical boxes quantify the typical scale of quantum fluctuations around $\langle V_3(\tau)\rangle$.

be neglected compared with the repulsive force due to the positive cosmological constant. Because the CDT simulations for technical reasons have to be performed in the Euclidean regime, we must compare the expectation value of the shape with those of the analytically continued expression of (13.4), with respect to the Euclidean time $\tau := -it$. After normalizing the overall four-volume and adjusting computer proper time by a constant to match continuum proper time, the averaged volume profile is depicted in Figure 13.4.

A few more things are noteworthy about this result. Firstly, despite the fact that the CDT construction deliberately breaks the isotropy between space and time, at least on large scales the full isotropy is restored by the ground state of the theory for precisely one choice of identifying proper time, that is, of fixing a relative scale between time and spatial distances in the continuum. Secondly, the computer simulations by necessity have to be performed for finite, compact spacetimes, which also means that a specific choice has to be made for the spacetime topology. For simplicity, to avoid having to specify boundary conditions, it is usually chosen to be $S^1 \times S^3$, with time compactified[5] and spatial slices which are topological three-spheres. What is reassuring is the fact that the bias this could in principle have introduced is 'corrected' by the system, which clearly is driven dynamically to the

[5] The period is chosen much larger than the time extension of the universe and does not influence the result.

topology of a four-sphere (as close to it as allowed by the kinematical constraint imposed on the three-volume, which is not allowed to vanish at any time). Lastly, we have also analysed the quantum fluctuations around the de Sitter background – they match to good accuracy a continuum saddlepoint calculation in minisuperspace [13], which is one more indication that we are indeed on the right track.

13.6 CDT key achievements III – a window on Planckian dynamics

Having discussed some of the evidence for obtaining the correct classical limit in CDT quantum gravity, let us turn to the *new* physics we are after, namely, what happens to gravity and the structure of spacetime at or near the Planck scale. We will describe one way of probing the short-scale structure, by setting up a *diffusion process* on the ensemble of spacetimes, and studying associated observables. The speed by which an initially localized diffusion process spreads into an ambient space is sensitive to the dimension of the space. Conversely, given a space M of unknown properties, it can be assigned a so-called *spectral dimension* D_S by studying the leading-order behaviour of the average return probability $\mathcal{R}_V(\sigma)$ (of random diffusion paths on M starting and ending at the same point x) as a function of the external diffusion time σ,

$$\mathcal{R}_V(\sigma) := \frac{1}{V(M)} \int_M d^d x \; P(x,x;\sigma) \propto \frac{1}{\sigma^{D_S/2}}, \quad \sigma \leq V^{2/D_S}, \tag{13.5}$$

where $V(M)$ is the volume of M, and $P(x_0, x; \sigma)$ the solution to the heat equation on M,

$$\partial_\sigma P(x_0, x, \sigma) = \nabla_x^2 P(x_0, x, \sigma). \tag{13.6}$$

Diffusion processes can be defined on very general spaces, for example, on fractals, which are partially characterized by their spectral dimension (usually not an integer, see [19]). Relevant for the application to quantum gravity is that the expectation value $\langle \mathcal{R}_V(\sigma) \rangle$ can be measured on the ensemble of CDT geometries, giving us the spectral dimension of the dynamically generated quantum universe, with the result that $D_S(\sigma)$ depends nontrivially on the diffusion time σ [8]! Since the linear scale probed in the diffusion is on the order of that of a random walker, $\sqrt{\sigma}$, short diffusion times probe the short-scale structure of geometry, and long ones its large-scale structure. The measurements from CDT quantum gravity, extrapolated to all values of σ, lead to the lower curve in Figure 13.5, with asymptotic values $D_S(0) = 1.82 \pm 0.25$, signalling highly nonclassical behaviour near the Planck scale, and $D_S(\infty) = 4.02 \pm 0.1$, which is compatible with the expected classical behaviour. Previous Euclidean models never showed such a scale dependence,

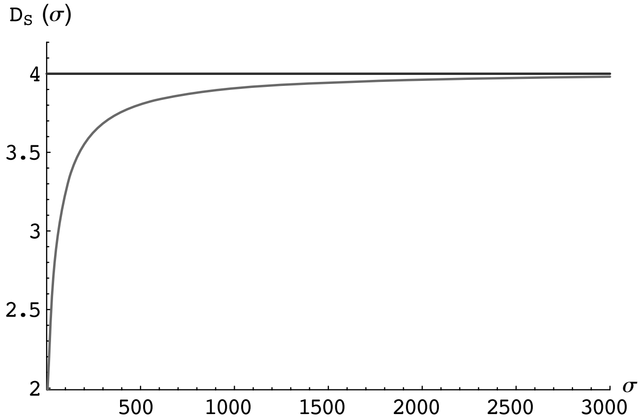


Figure 13.5 The spectral dimension $D_S(\sigma)$ of the CDT-generated quantum universe (lower curve, error bars not included), contrasted with the corresponding curve for a classical spacetime, simply given by the constant function $D_S(\sigma) = 4$. We assume $\sigma \ll V^{2/D_S}$, so no finite-volume effects are present.

reflecting their lack of an interesting geometric structure as a function of scale. For CDT in three spacetime dimensions, there is evidence for an analogous scale dependence [20].

This somewhat unexpected result found in nonperturbative CDT quantum gravity has brought into focus the role of 'dynamical dimensions' as diffeomorphism-invariant indicators of nonclassicality at the Planck scale.[6] Interestingly, a similar dimensional reduction from four to two near the Planck scale has since been found in disparate approaches to quantum gravity, most prominently, a nonperturbative renormalization group flow analysis [33], and so-called Lifshitz gravity [29]. The coincidence is certainly intriguing and could mean that at a more fundamental level the approaches have more in common than we currently understand (*and* capture a true aspect of nonperturbative quantum gravity). Reproducing dimensional reduction could then even become a 'test' of quantum gravity, similar to how the derivation of the black hole entropy formula $S_{\mathrm{BH}} = A/4$ is often viewed, with the difference that the latter is usually associated with a semiclassical context, whereas the former is thought to characterize the behaviour of the theory in the deep UV.

Further evidence of nonclassicality on short scales in CDT comes from measurements of geometric structures in spatial slices $\tau = \text{const.}$, including a measurement of their Hausdorff and spectral dimensions [9].

[6] Other notions of dimensionality are the Hausdorff and the fractal dimension.

13.7 Open issues and outlook

As we have summarized above, significant strides have been made in the causal dynamical triangulations quantization program in demonstrating its compatibility with classical general relativity on large scales, and at the same time exploring its true quantum properties on small scales. A number of important issues are the subject of ongoing and future research. Firstly, as explained in more detail in [13], one would like to tune the bare parameters of the CDT simulations so as to obtain a better length resolution and get even closer to and, if possible, below the Planck scale. (Current simulations operate with quantum universes of the order of 10–20 Planck lengths across.) This would enable us to look for more direct evidence of the existence or otherwise of the nontrivial UV fixed point seen in truncated renormalization group flows [32].

An important challenge is to reproduce further aspects of the classical limit correctly, one of which is the derivation of Newton's law 'from scratch' in the nonperturbative theory. As a possibly first step towards this goal, some physical consequences of the presence of an isolated point mass in CDT's quantum de Sitter universe have been analyzed in [30]. Another natural area of application is the early universe, with or without the addition of a scalar field ('inflaton'), to check and discriminate between the often ambiguous and contradictory claims of quantum cosmological models in a context where *all* fluctuations of the geometry are present, not just the overall scale factor. A concrete example of what one might be able to do is nailing down factor-ordering ambiguities in the cosmological path integral [37].

Coming back to some of the questions we raised at the outset of this chapter, the preliminary conclusion about the nature of quantum spacetime is that it is nothing like a four-dimensional classical manifold on short scales. In addition to its anomalous spectral dimension, its naive Regge curvature diverges, indicating a singular behaviour reminiscent of (but surely worse than) that of the particle paths constituting the support of the Wiener measure. However, it apparently is *not* literally a 'spacetime foam', if by that one means some bubbling, topology-changing entity: one of the main findings of dynamically triangulated models of nonperturbative quantum gravity is that allowing for local topology changes and making them part of the dynamics renders the quantum superposition inherently unstable and is incompatible with a good classical limit. Even if local topology change is not part of quantum-gravitational dynamics, we saw that global topology, as well as short-scale dimensionality, are determined dynamically and do not necessarily coincide with the (somewhat arbitrary) choices made for them as part of the regularized formulation. What makes these perhaps surprising findings possible is the fact that CDT quantum gravity allows for large curvature fluctuations on short scales, and that the construction of the final theory involves a nontrivial limiting process, which the computer simulations are able to approximate.

In summary, if there is indeed a unique, interacting quantum field theory of spacetime geometry in four dimensions, which does not contain any exotic ingredients, and has general relativity as its classical limit, the CDT approach has a good chance of finding it. It relies only on a minimal set of ingredients and priors: the quantum superposition principle, locality (micro-)causality, a notion of (proper) time and standard tools from quantum field theory otherwise,[7] has few free parameters (essentially the couplings of the phase diagram of Figure 13.3), and by virtue of its construction through a scaling limit can rely on a considerable degree of universality in the sense of critical system theory. Although many issues remain to be tackled and understood, the interesting new results and insights CDT has produced to date make for a pretty good start.

Acknowledgments

All the authors gratefully acknowledge support by ENRAGE (European Network on Random Geometry), a Marie Curie Research Training Network, contract MRTN-CT-2004-005616. In addition, JJ was supported by COCOS (Correlations in Complex Systems), a Marie Curie Transfer of Knowledge Project, contract MTKD-CT-2004-517186, and RL by the Netherlands Organisation for Scientific Research (NWO) under their VICI program.

References

[1] J. Ambjørn and R. Loll: Non-perturbative Lorentzian quantum gravity, causality and topology change, *Nucl. Phys. B* **536** (1998) 407–34 [hep-th/9805108].

[2] J. Ambjørn and Yu. M. Makeenko: Properties of loop equations for the Hermitean matrix model and for two-dimensional quantum gravity, *Mod. Phys. Lett. A* **5** (1990) 1753–64; J. Ambjørn, J. Jurkiewicz and Y. M. Makeenko: Multiloop correlators for two-dimensional quantum gravity, *Phys. Lett. B* **251** (1990) 517–24.

[3] 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]; Crossing the $c = 1$ barrier in 2d Lorentzian quantum gravity, *Phys. Rev. D* **61** (2000) 044010 [hep-lat/9909129].

[4] J. Ambjørn, J. Jurkiewicz and R. Loll: A nonperturbative Lorentzian path integral for gravity, *Phys. Rev. Lett.* **85** (2000) 924–7 [hep-th/0002050]; Dynamically triangulating Lorentzian quantum gravity, *Nucl. Phys. B* **610** (2001) 347–82 [hep-th/0105267].

[5] J. Ambjørn, J. Jurkiewicz and R. Loll: Lorentzian and Euclidean quantum gravity: Analytical and numerical results, in: *Proceedings of M-Theory and Quantum Geometry*, 1999 NATO Advanced Study Institute, Akureyri Island, eds. L. Thorlacius *et al.* (Kluwer, 2000) 382–449 [hep-th/0001124].

[6] J. Ambjørn, J. Jurkiewicz and R. Loll: Nonperturbative 3-d Lorentzian quantum gravity, *Phys. Rev. D* **64** (2001) 044011 [hep-th/0011276].

[7] This puts it about on a par with the renormalization group approach of [39, 40], makes fewer assumptions than loop quantum gravity [44], but is not quite as minimalistic as the causal set approach [27].

[7] J. Ambjørn, J. Jurkiewicz and R. Loll: Emergence of a 4D world from causal quantum gravity, *Phys. Rev. Lett.* **93** (2004) 131301 [hep-th/0404156].
[8] J. Ambjørn, J. Jurkiewicz and R. Loll: Spectral dimension of the universe, *Phys. Rev. Lett.* **95** (2005) 171301 [hep-th/0505113].
[9] J. Ambjørn, J. Jurkiewicz and R. Loll: Reconstructing the universe, *Phys. Rev. D* **72** (2005) 064014 [hep-th/0505154].
[10] J. Ambjørn, A. Görlich, J. Jurkiewicz and R. Loll: Planckian birth of the quantum de Sitter universe, *Phys. Rev. Lett.* **100** (2008) 091304 [arXiv:0712.2485, hep-th].
[11] J. Ambjørn, R. Loll, Y. Watabiki, W. Westra and S. Zohren: A string field theory based on Causal Dynamical Triangulations, *JHEP* **0805** (2008) 032 [arXiv:0802.0719, hep-th].
[12] J. Ambjørn, K. N. Anagnostopoulos, R. Loll and I. Pushkina: Shaken, but not stirred – Potts model coupled to quantum gravity, *Nucl. Phys. B* **807** (2009) 251 [arXiv:0806.3506, hep-lat].
[13] J. Ambjørn, A. Görlich, J. Jurkiewicz and R. Loll: The nonperturbative quantum de Sitter universe, *Phys. Rev. D* **78** (2008) 063544 [arXiv:0807.4481, hep-th].
[14] J. Ambjørn, J. Jurkiewicz and R. Loll: The self-organizing quantum universe, *Sci. Am.* **299**N1 (2008) 42–9; The self-organized de Sitter universe, *Int. J. Mod. Phys. D* **17** (2009) 2515–20 [arXiv:0806.0397, gr-qc].
[15] J. Ambjørn, J. Jurkiewicz and R. Loll: Quantum gravity as sum over spacetimes [arXiv:0906.3947, gr-qc]; Quantum gravity, or the art of building spacetime, in: *Approaches to Quantum Gravity*, ed. D. Oriti, Cambridge University Press (2009) 341–59 [hep-th/0604212]; R. Loll: The emergence of spacetime, or, Quantum gravity on your desktop, *Class. Quant. Grav.* **25** (2008) 114006 [arXiv:0711.0273, gr-qc]; J. Ambjørn, A. Görlich, J. Jurkiewicz and R. Loll: The quantum universe, *Acta Phys. Polon. B* **39** (2008) 3309–41.
[16] J. Ambjørn, R. Loll, W. Westra and S. Zohren; Summing over all topologies in CDT string field theory, *Phys. Lett. B* **678** (2009) 227 [arXiv:0905.2108, hep-th].
[17] J. Ambjørn, R. Loll, Y. Watabiki, W. Westra and S. Zohren: Proper time is stochastic time in 2d quantum gravity [arXiv:0911.4211, hep-th].
[18] B. V. de Bakker: Further evidence that the transition of 4D dynamical triangulation is 1st order, *Phys. Lett. B* **389** (1996) 238 [hep-lat/9603024].
[19] D. ben-Avraham and S. Havlin: Diffusion and Reactions in Fractals and Disordered Systems, Cambridge University Press (2000).
[20] D. Benedetti and J. Henson: Spectral geometry as a probe of quantum spacetime, *Phys. Rev. D* **80** (2009) 124036 [arXiv:0911.0401, hep-th].
[21] P. Bialas, Z. Burda, A. Krzywicki and B. Petersson: Focusing on the fixed point of 4d simplicial gravity, *Nucl. Phys. B* **472** (1996) 293 [hep-lat/9601024].
[22] P. Bialas, Z. Burda, B. Petersson and J. Tabaczek: Appearance of mother universe and singular vertices in random geometries, *Nucl. Phys. B* **495** (1997) 463 [hep-lat/9608030].
[23] G. F. R. Ellis: Issues in the philosophy of cosmology [astro-ph/0602280].
[24] G. W. Gibbons and S. W. Hawking (eds): *Euclidean Quantum Gravity*, World Scientific, Singapore (1993).
[25] S. B. Giddings: Nonlocality vs. complementarity: a conservative approach to the information problem [arXiv:0911.3395, hep-th].
[26] J. J. Halliwell and J. Louko: Steepest descent contours in the path integral approach to quantum cosmology. 3. A general method with applications to anisotropic minisuperspace models, *Phys. Rev. D* **42** (1990) 3997–4031.

[27] J. Henson: The causal set approach to quantum gravity, in: *Approaches to Quantum Gravity*, ed. D. Oriti, Cambridge University Press (2009) 393-413 [gr-qc/0601121].
[28] G. 't Hooft: Emergent quantum mechanics and emergent symmetries, *AIP Conf. Proc.* **957** (2007) 154 [arXiv:0707.4568, hep-th].
[29] P. Hořava: Spectral dimension of the universe in quantum gravity at a Lifshitz point, *Phys. Rev. Lett.* **102** (2009) 161301 [arXiv:0902.3657, hep-th].
[30] I. Khavkine, R. Loll and P. Reska: Coupling point-like masses to quantum gravity with causal dynamical triangulations, preprint Utrecht U., *Class Quant. Grav.* **27** (2010) 185025 [arXiv:1002.4618, gr-qc].
[31] C. Kiefer: *Quantum Gravity*, 2nd edn, Oxford University Press (2007).
[32] O. Lauscher and M. Reuter: Ultraviolet fixed point and generalized flow equation of quantum gravity, *Phys. Rev. D* **65** (2002) 025013 [hep-th/0108040].
[33] O. Lauscher and M. Reuter: Fractal spacetime structure in asymptotically safe gravity, *JHEP* **0510** (2005) 050 [hep-th/0508202].
[34] R. Loll: The volume operator in discretized quantum gravity, *Phys. Rev. Lett.* **75** (1995) 3048 [gr-qc/9506014].
[35] R. Loll and W. Westra: Sum over topologies and double-scaling limit in 2D Lorentzian quantum gravity, *Class. Quant. Grav.* **23** (2006) 465 [hep-th/0306183].
[36] R. Loll, W. Westra and S. Zohren: Taming the cosmological constant in 2D causal quantum gravity with topology change, *Nucl. Phys. B* **751** (2006) 419 [hep-th/0507012].
[37] R. L. Maitra: Can causal dynamical triangulations probe factor-ordering issues?, *Acta Phys. Polon. B Proc. Suppl.* **2** (2009) 563 [arXiv:0910.2117, gr-qc].
[38] M. E. J. Newman and G. T. Barkema: *Monte Carlo Methods in Statistical Physics*, Clarendon Press, Oxford (1999).
[39] M. Niedermaier: The asymptotic safety scenario in quantum gravity: An introduction, *Class. Quant. Grav.* **24** (2007) R171 [gr-qc/0610018].
[40] M. Niedermaier and M. Reuter: The asymptotic safety scenario in quantum gravity, *Living Rev. Rel.* **9** (2006) 5.
[41] M. Reed and B. Simon: *Methods of Modern Mathematical Physics, vol. 2*, Academic Press (1975).
[42] T. Regge: General relativity without coordinates, *Nuovo Cim. A* **19** (1961) 558–71.
[43] C. Rovelli and L. Smolin: Discreteness of area and volume in quantum gravity, *Nucl. Phys. B* **442** (1995) 593; Erratum-ibid. **456** (1995) 753 [gr-qc/9411005].
[44] T. Thiemann: Loop quantum gravity: An inside view, *Lect. Notes Phys.* **721** (2007) 185 [hep-th/0608210].

14

Proper time is stochastic time in 2D quantum gravity

J. AMBJØRN, R. LOLL, Y. WATABIKI,
W. WESTRA & S. ZOHREN

We show that proper time, when defined in the quantum theory of 2D gravity, becomes identical to the stochastic time associated with the stochastic quantization of space. This observation was first made by Kawai and collaborators in the context of 2D Euclidean quantum gravity, but the relation is even simpler and more transparent in the context of 2D gravity formulated in the framework of CDT (causal dynamical triangulations).

14.1 Introduction

Since time plays such a prominent role in ordinary flat space quantum field theory defined by a Hamiltonian, it is of interest to study the role of time even in toy models of quantum gravity where the role of time is much more enigmatic. The model we will describe in this chapter is the so-called causal dynamical triangulation (CDT) model of quantum gravity. It starts by providing an ultraviolet regularization in the form of a lattice theory, the lattice link length being the (diffeomorphism-invariant) UV cut-off. In addition, the lattice respects causality. It is formulated in the spirit of *asymptotic safety*, where it is assumed that quantum gravity is described entirely by "conventional" quantum field theory, in this case by approaching a non-trivial fixed point [1, 2]. It is formulated in space-times with Lorentzian signature, but the regularized space-times which are used in the path integral defining the theory allow a rotation to Euclidean space-time. The action used is the Regge action for the piecewise linear geometry. Each (piecewise linear) geometry used in the path integral has after rotation to Euclidean signature a Euclidean Regge action, related to the original Lorentzian action in the same way as when one in flat space-time

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

rotates Lorentzian time to Euclidean time (see [3, 4] for details of the Regge action in the CDT approach). By rotating each Lorentzian CDT lattice (or piecewise linear geometry) to Euclidean signature, the non-perturbative path integral is performed by summing over a set of Euclidean lattices originating from the Lorentzian lattices with a causal stucture, and this set is different from the full set of piecewise linear Euclidean lattices. Like in ordinary lattice field theories, we approach the continuum by fine-tuning the bare coupling constants. The rotation to Euclidean space-time makes it possible to use Monte Carlo simulations when studying the theory, and in four-dimensional space-time, which for obvious reasons has our main interest, there exists a region of coupling constant space where the infrared behavior of the universe seen by the computer is that of (Euclidean) de Sitter space-time [5, 6] (for a pedagodical review, see [7]). Non-trivial UV properties have been observed [8], properties which have been reproduced by other "field theoretical" approaches to quantum gravity [9, 10].

Numerical simulations are very useful when trying to understand if a non-perturbatively defined quantum field theory has a chance to make sense. However, numerical simulations have their limitations in the sense that they will never provide a proof of the existence of a theory and it might be difficult in detail to follow the way the continuum limit is approached since it requires larger and larger lattices. It is thus of interest and importance to be able to study this in detail, even if only in a toy model. Two-dimensional quantum gravity is such a toy model which has a surprisingly rich structure. Many of the intriguing questions in quantum gravity and in lattice quantum gravity are still present in the two-dimensional theory. We will discuss the solution to two-dimensional CDT in the rest of this chapter and we will see that *time*, which is introduced as *proper time*, has an interpretation as *stochastic time* in a process where the evolution of space can be viewed as stochastic.

14.2 The CDT formalism

We start from Lorentzian simplicial space-times with $d = 2$ and insist that only causally well-behaved geometries appear in the (regularized) Lorentzian path integral. A crucial property of our explicit construction is that each of the configurations allows for a rotation to Euclidean signature, as mentioned above. We rotate to a Euclidean regime in order to perform the sum over geometries (and rotate back again afterward if needed). We stress here (again) that although the sum is performed over geometries with Euclidean signature, it is different from what one would obtain in a theory of quantum gravity based *ab initio* on Euclidean space-times. The reason is that not all Euclidean geometries with a given topology are included in the "causal" sum since in general they have no correspondence to a causal Lorentzian geometry.

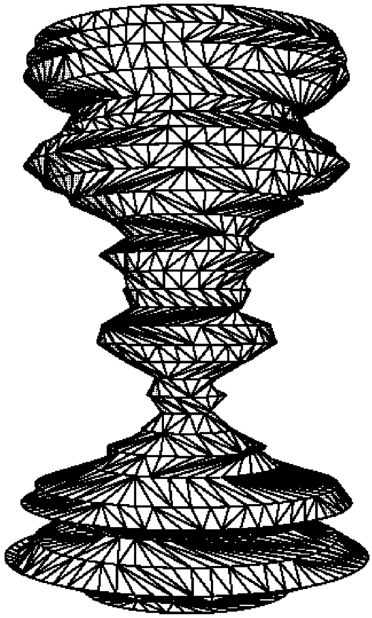


Figure 14.1 Piecewise linear space-time histories $(1+1)$ dimensional quantum gravity.

We refer to [3] for a detailed description of how to construct the class of piecewise linear geometries used in the Lorentzian path integral in higher dimensions. The most important assumption is the existence of a global proper-time foliation. This is illustrated in Figure 14.1 in the case of two dimensions. We have a sum over two-geometries, "stretching" between two "one-geometries" separated by a proper time t and constructed from two-dimensional building blocks. In Figure 14.2 we have shown how to fill the two-dimensional space-time between the space (with topology S^1) at time t_n and time $t_{n+1} = t_n + a$, where a denotes the lattice spacing. While in the lattice model we often use units where everything is measured in lattice lengths (i.e. the lattice links have length one), we are of course interested in taking the limit $a \to 0$ to recover continuum physics.

In the path integral we will be summing over all possible ways to connect a given 1D "triangulation" at time t_n and a given 1D triangulation at time t_{n+1} to a slab of 2D space-time as shown in Figure 14.2, and in addition we will sum over all 1D "triangulations" of S^1 at times t_n. Thus we are demanding that the time-slicing is such that the topology of space does not change when space "evolves" from time t_n to time t_{n+1}.

The Einstein–Hilbert action S^{EH} in two dimensions is trivial since there is no curvature term as long as the topology of space-time is unchanged (which we assume

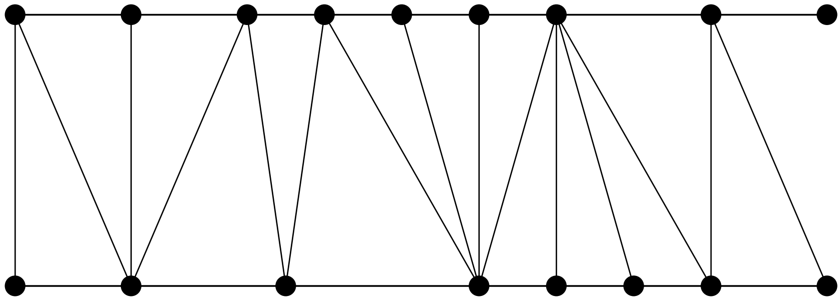


Figure 14.2 The propagation of a spatial slice from time t to time $t+1$. The end of the strip should be joined to form a band with topology $S^1 \times [0,1]$.

presently). Thus the (Euclidean) action simply consists of the cosmological term:

$$S_E^{\mathrm{EH}} = \lambda \int d^2x \sqrt{g} \quad \longrightarrow \quad S_E^{\mathrm{Regge}} = \Lambda N_2, \tag{14.1}$$

where N_2 denotes the total number of triangles in the two-dimensional triangulation. We denote the discretized action as *the Regge action*, since it is a trivial example of the natural action for piecewise linear geometries introduced by Regge [11]. The dimensionless lattice cosmological coupling constant Λ will be related to the continuum cosmological coupling constant λ by an additive renormalization:

$$\Lambda = \Lambda_0 + \frac{1}{2}\lambda a^2, \tag{14.2}$$

the factor 1/2 being conventional. The path integral or partition function for the CDT version of quantum gravity is now

$$G_\lambda^{(0)}(l_1, l_2; t) = \int \mathcal{D}[g]\, e^{-S_E^{\mathrm{EH}}[g]} \quad \rightarrow$$

$$G_\Lambda^{(0)}(L_1, L_2, T) = \sum_{\mathcal{T}} \frac{1}{C_{\mathcal{T}}}\, e^{-S_E^{\mathrm{Regge}}(\mathcal{T})}, \tag{14.3}$$

where the summation is over all causal triangulations $\mathcal{T}$ of the kind described above with a total of T time steps, an "entrance loop" of length $l_1 = L_1 a$, and an "exit loop" of length $l_2 = L_2 a$. The factor $1/C_{\mathcal{T}}$ is a symmetry factor, given by the order of the automorphism group of the triangulation $\mathcal{T}$.

One can, somewhat surprisingly, evaluate the sum over triangulations in (14.3) analytically [12]. It is a counting problem and thus we introduce the corresponding generating function. In our model the generating function has a direct physical interpretation. We define

$$\widetilde{G}_{\Lambda}^{(0)}(X_1, X_2; t) = \sum_{L_1, L_2} \mathrm{e}^{-X_1 L_1} \mathrm{e}^{-X_2 L_2} G_{\Lambda}^{(0)}(L_1, L_2; T). \tag{14.4}$$

Thus $\widetilde{G}_{\Lambda}^{(0)}(X_1, X_2; T)$ is the generating function of the numbers $G_{\Lambda}^{(0)}(L_1, L_2; T)$ if we write $Z_1 = e^{-X_1}$, $Z_2 = e^{-X_2}$. But we can also view X as a (bare) dimensionless boundary cosmological constant, such that a boundary cosmological term $X \cdot L$ has been added to the action. In this way $G_{\Lambda}^{(0)}(X_1, X_2; T)$ represents the sum over triangulations where the lengths of the boundaries are allowed to fluctuate, the fluctuations controlled by the values X_i of the boundary cosmological constants. In general we expect, just based on standard dimensional analysis, the boundary cosmological constants X_i to be subject to additive renormalizations when the continuum limit is approached. Like (14.2) we expect

$$X_i = X_c + x_i\, a, \tag{14.5}$$

where x then denotes the continuum boundary cosmological constant, and one, after renormalization, has the continuum boundary cosmological action $x \cdot l$.

We refer to [12] for the explicit combinatorial arguments which allow us to find $\widetilde{G}_{\Lambda}^{(0)}(X_1, X_2; T)$. Let us just state the following results: one can derive an exact iterative equation (using notation $Z = \mathrm{e}^{-X}$, $W = \mathrm{e}^{-Y}$, $Q = \mathrm{e}^{-\Lambda}$)

$$\widetilde{G}_{\Lambda}^{(0)}(Z, W; T) = \frac{QZ}{1 - QZ}\, \widetilde{G}_{\Lambda}^{(0)}\Big(\frac{Q}{1 - QZ}, W; T - 1\Big). \tag{14.6}$$

This equation can be iterated and the solution found [12]. However, it is easy to see that $Q_c = 1/2$ and that $Z_c = 1$ and we can now take the continuum limit in (14.6) using $t = T \cdot a$ and find

$$\frac{\partial}{\partial t} \widetilde{G}_{\lambda}^{(0)}(x, y; t) + \frac{\partial}{\partial x}\Big[(x^2 - \lambda)\widetilde{G}_{\lambda}^{(0)}(x, y; t)\Big] = 0. \tag{14.7}$$

This is a standard first-order partial differential equation which should be solved with the boundary condition

$$\widetilde{G}_{\lambda}^{(0)}(x, y; t = 0) = \frac{1}{x + y} \tag{14.8}$$

corresponding to

$$G_{\lambda}^{(0)}(l_1, l_2; t = 0) = \delta(l_1 - l_2). \tag{14.9}$$

The solution is thus

$$\widetilde{G}_{\lambda}^{(0)}(x, y; t) = \frac{\bar{x}^2(t; x) - \lambda}{x^2 - \lambda}\, \frac{1}{\bar{x}(t; x) + y}, \tag{14.10}$$

where $\bar{x}(t;x)$ is the solution to the characteristic equation

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = -(\bar{x}^2 - \lambda), \quad \bar{x}(t=0) = x. \tag{14.11}$$

We thus have an explicit solution for $\widetilde{G}_\lambda^{(0)}(x,y;t)$, since we obtain

$$\bar{x}(t) = \sqrt{\lambda}\, \frac{(\sqrt{\lambda}+x) - \mathrm{e}^{-2\sqrt{\lambda}t}(\sqrt{\lambda}-x)}{(\sqrt{\lambda}+x) + \mathrm{e}^{-2\sqrt{\lambda}t}(\sqrt{\lambda}-x)}. \tag{14.12}$$

If we interpret the propagator $G_\lambda^{(0)}(l_1,l_2;t)$ as the matrix element between two boundary states of a Hamiltonian evolution in "time" t,

$$G_\lambda^{(0)}(l_1,l_2;t) = < l_1|\mathrm{e}^{-H_0 t}|l_2 > \tag{14.13}$$

we can, after an inverse Laplace transformation, read off the functional form of the Hamiltonian operator H_0 from (14.7):

$$\widetilde{H}_0(x) = \frac{\partial}{\partial x}\Big(x^2 - \lambda\Big), \quad H_0(l) = -l\frac{\partial^2}{\partial l^2} + \lambda l. \tag{14.14}$$

This ends our short review of basic 2D CDT. We have emphasized here that all continuum results can be obtained by solving the lattice model explicitly and taking the continuum limit simply by letting the lattice spacing $a \to 0$. The same will be true for the generalized CDT model described below, but to make the presentation more streamlined we will drop the explicit route via a lattice and work directly in the continuum.

14.3 Generalized CDT

It is natural to ask what happens if the strict requirement of "classical" causality on each geometry appearing in the path integral is relaxed. While causality is a reasonable requirement as an outcome of a sensible physical theory, there is no compelling reason to impose it on each individual geometry in the path integral, since these are not physical observables. We used it, inspired by [13], as a guiding principle for obtaining a path integral which is different from the standard Euclidean path integral, which was seemingly a necessity in higher than two space-time dimensions since a "purely" Euclidean higher-dimensional path integral did not lead to interesting continuum theories.

In Figure 14.3 we show what happens if we allow causality to be violated locally by allowing space to split in two at a certain time t, but we never allow the "baby" universe which splits off to come back to the "parent" universe. The baby universe

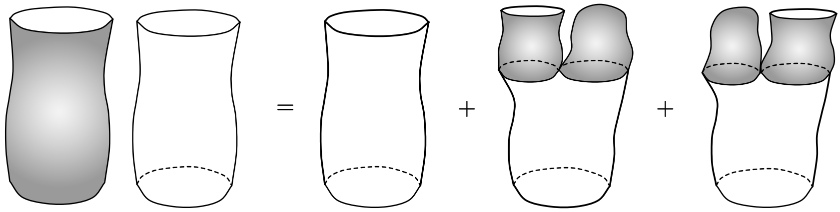


Figure 14.3 In all four graphs, the geodesic distance from the final to the initial loop is given by t. Differentiating with respect to t leads to Eq. (14.15). Shaded parts of the graphs represent the full, g_s-dependent propagator and disk amplitude, and non-shaded parts the CDT propagator.

thus continues its life and is assumed to vanish, shrink to nothing, at some later time. We now integrate over all such configurations in the path integral. From the point of view of Euclidean space-time, we are simply integrating over all space-times with the topology of a cylinder. However, returning to the original Lorentzian picture it is clear that at the point where space splits in two the light-cone is degenerate and one is violating causality in the strict local sense that each space-time point should have a future and a past light-cone. Similarly, when the baby universe "ends" its time evolution the light-cone structure is degenerate. These points thus have a diffeomorphism-invariant meaning in space-times with Lorentzian structure, and it makes sense to associate a coupling constant g_s with the process of space branching in two disconnected pieces.

The equation corresponding to Figure 14.3 is [14]

$$\frac{\partial}{\partial t}\widetilde{G}_{\lambda,g_s}(x,y;t) = -\frac{\partial}{\partial x}\Big[\Big((x^2-\lambda)+2g_s\,\widetilde{W}_{\lambda,g_s}(x)\Big)\widetilde{G}_{\lambda,g_s}(x,y;t)\Big]. \tag{14.15}$$

$\widetilde{W}_{\lambda,g_s}(x)$ denotes the disk amplitude with a fixed boundary cosmological constant x. It is related to the disk amplitude with a fixed boundary length by

$$\widetilde{W}_{\lambda,g_s}(x) = \int_0^\infty \mathrm{d}l\,\mathrm{e}^{-xl}\,W_{\lambda,g_s}(l). \tag{14.16}$$

It describes the "propagation" of a spatial universe until it vanishes in the vacuum. If we did not allow any spatial branching we would simply have

$$\widetilde{W}^{(0)}_\lambda(x) = \int_0^\infty \mathrm{d}t\; G^{(0)}_\lambda(x,l=0;t) = \frac{1}{x+\sqrt{\lambda}}, \tag{14.17}$$

where $G^{(0)}_\lambda(x,l;t)$ denotes the Laplace transform of $G^{(0)}_\lambda(l',l;t)$ with respect to l'. From the composition rules for $G_{\lambda,g_s}(l_1,l_2;t)$ it follows that it has (mass) dimension 1. Thus $G_{\lambda,g_s}(x,l_2;t)$ is dimensionless and it follows that the (mass) dimension of the coupling constant g_s must be 3. In a discretized theory it will appear as the dimensionless combination $g_s a^3$, a being the lattice spacing, and one can show that the creation of more than one baby universe at a given time t is suppressed by powers of a (see [14] for details). Thus we only need to consider the process shown in Figure 14.3. For a fixed cosmological constant λ and boundary cosmological constants x, y, expressions like $\widetilde{G}_{\lambda,g_s}(x, y; t)$ and $\widetilde{W}_{\lambda,g_s}(x)$ will have a power series expansion in the dimensionless variable

$$\kappa = \frac{g_s}{\lambda^{3/2}} \tag{14.18}$$

and the radius of convergence is of order one. Thus the coupling constant g_s indeed acts to tame the creation of baby universes and if g_s exceeds this critical value Eq. (14.15) breaks down and is replaced by another equation corresponding to Liouville quantum gravity with central change $c = 0$ (see [14] for a detailed discussion).

Differentiating the integral equation corresponding to Figure 14.3 with respect to the time t, one obtains (14.15). The disk amplitude $\widetilde{W}_{\lambda,g_s}(x)$ is at this stage unknown. However, one has the graphical representation for the disk amplitude shown in Figure 14.4. It translates into the equation [14]

$$\begin{aligned}\widetilde{W}_{\lambda,g_s}(x) = \widetilde{W}^{(0)}_{\lambda}(x) + g_s \int_0^\infty \mathrm{d}t \int_0^\infty \mathrm{d}l_1 \mathrm{d}l_2\ (l_1 + l_2) G^{(0)}_{\lambda}(x, l_1 + l_2; t)& \\ \times W_{\lambda,g_s}(l_1) W_{\lambda,g_s}(l_2).&\end{aligned} \tag{14.19}$$

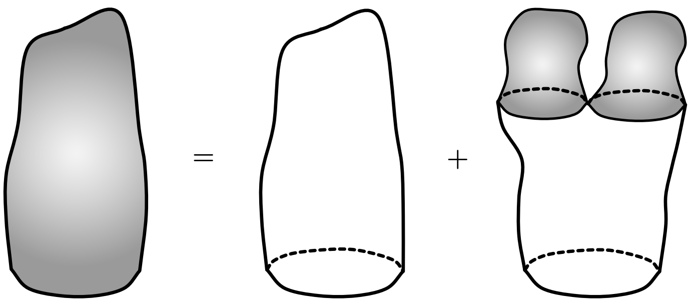


Figure 14.4 Graphical illustration of Eq. (14.19). Shaded parts represent the full disk amplitude, unshaded parts the CDT disk amplitude and the CDT propagator.

The superscript (0) indicates the CDT amplitudes without baby universe branching, calculated above. We assume

$$\widetilde{W}_{\lambda,g_s=0}(x) = \widetilde{W}^{(0)}_{\lambda}(x), \tag{14.20}$$

and similarly for $G^{(0)}_{\lambda,g_s}$. The integrations in (14.19) can be performed and if we define $\hat{W}_{\lambda,g_s}(x)$ by

$$\widetilde{W}_{\lambda,g_s}(x) = \frac{-(x^2-\lambda)+\hat{W}_{\lambda,g_s}(x)}{2g_s}, \tag{14.21}$$

one can show that $\hat{W}_{\lambda,g_s}(x)$ is given by the expression

$$\hat{W}_{\lambda,g_s}(x) = \lambda(\widetilde{x}-u)\sqrt{(\widetilde{x}+u)^2-2\kappa}, \tag{14.22}$$

where

$$x = \widetilde{x}\sqrt{\lambda}, \quad u^3-u+\kappa = 0. \tag{14.23}$$

In order to have a physically acceptable $\widetilde{W}_{\lambda,g_s}(x)$, one has to choose the solution to the third-order equation which is closest to 1 and the above statements about the expansion of $\widetilde{W}_{\lambda,g_s}(x)$ in a power series in κ follow.

Equation (14.15) can now be written as

$$\frac{\partial}{\partial t}\widetilde{G}_{\lambda,g_s}(x,y;t) = -\frac{\partial}{\partial x}\Big[\hat{W}_{\lambda,g_s}(x)\,\widetilde{G}_{\lambda,g_s}(x,y;t)\Big]. \tag{14.24}$$

In analogy with (14.7) and (14.10), this is solved by

$$\widetilde{G}_{\lambda,g_s}(x,y;t) = \frac{\hat{W}_{\lambda,g_s}(\bar{x}(t,x))}{\hat{W}_{\lambda,g_s}(x)}\,\frac{1}{\bar{x}(t,x)+y}, \tag{14.25}$$

where $\bar{x}(t,x)$ is the solution of the characteristic equation for (14.24), the generalization of Eq. (14.11):

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = -\hat{W}_{\lambda,g_s}(\bar{x}), \quad \bar{x}(0,x) = x, \tag{14.26}$$

such that

$$t = \int_{\bar{x}(t)}^{x} \frac{\mathrm{d}y}{\hat{W}_{\lambda,g_s}(y)}. \tag{14.27}$$

This integral can be expressed in terms of elementary functions and one can thus find an explicit expression for $\widetilde{G}_{\lambda,g_s}(x,y;t)$ in the same way as Eq. (14.12) led to an explicit solution for the $\widetilde{G}^{(0)}_{\lambda}(x,y;t)$ appearing in (14.10).

14.4 The matrix model representation

The formulas (14.22) and (14.21) are standard formulas for the resolvent of a hermitian matrix model, calculated to leading order in N, the size of the matrix. In fact, the following matrix model:

$$Z(\lambda, g_s) = \int d\phi \, \mathrm{e}^{-N \mathrm{Tr}\, V(\phi)}, \qquad V(\phi) = \frac{1}{g_s}\Big(\lambda\phi - \frac{1}{3}\phi^3\Big) \tag{14.28}$$

has resolvent

$$\left\langle \frac{1}{N} \mathrm{Tr} \left(\frac{1}{x-\phi} \right) \right\rangle = \widetilde{W}_{\lambda, g_s(x)}(x) + O(1/N^2), \tag{14.29}$$

where $\widetilde{W}_{\lambda, g_s}(x)$ is given by (14.21), and where the expectation value of a matrix expression $\mathcal{O}(\phi)$ is defined as

$$\langle \mathcal{O}(\phi) \rangle = \frac{1}{Z(\lambda, g_s)} \int d\phi \, \mathrm{e}^{-N \mathrm{Tr}\, V(\phi)} \, \mathcal{O}(\phi). \tag{14.30}$$

What is surprising here, compared to the "old" matrix model approaches to 2D Euclidean quantum gravity, is that the large-N limit reproduces directly the continuum theory. No scaling limit has to be taken. The situation is more like in the Kontsevich matrix model, which describes continuum 2D gravity aspects directly. In fact, the cubic potential is "almost" like the cubic potential in the Kontsevich matrix model, but the worldsheet interpretation is different.

Can the above correspondence be made systematic in a large-N expansion and can the matrix model representation help us to a non-perturbative definition of generalized 2D CDT gravity? The answer is yes [15].

First we have to formulate the CDT model *from first principles* such that we allow for baby universes to join the "parent" universe again, i.e. we have to allow for topology changes of the 2D universe, and next we have to check if this generalization is correctly captured by the matrix model (14.28) [16].

14.5 CDT string field theory

In quantum field theory particles can be created and annihilated if the process does not violate any conservation law of the theory. In string field theories one operates in the same way with operators which can create and annihilate strings. From the 2D quantum gravity point of view we thus have a third-quantization of gravity: one-dimensional universes can be created and destroyed. In [17] such a formalism was developed for non-critical strings (or 2D Euclidean quantum gravity). In [16] the formalism was applied to 2D CDT gravity leading to a string field theory or third quantization for CDT, which allows us in principle to calculate any amplitude involving creation and annihilation of universes.

Let us briefly review this formalism. The starting point is the assumption of a vacuum from which universes can be created. We denote this state $|0\rangle$ and define creation and annihilation operators:

$$[\Psi(l), \Psi^\dagger(l')] = l\delta(l-l'), \quad \Psi(l)|0\rangle = \langle 0|\Psi^\dagger(l) = 0. \tag{14.31}$$

The factor l multiplying the delta-function is introduced for convenience, see [16] for a discussion.

Associated with the spatial universe we have a Hilbert space on the positive half-line, and a corresponding scalar product (making $H_0(l)$ defined in Eq. (14.14) hermitian):

$$\langle \psi_1|\psi_2\rangle = \int \frac{\mathrm{d}l}{l}\, \psi_1^*(l)\psi_2(l). \tag{14.32}$$

The introduction of the operators $\Psi(l)$ and $\Psi^\dagger(l)$ in (14.31) can be thought of as analogous to the standard second-quantization in many-body theory. The single-particle Hamiltonian H_0 defined by (14.14) becomes in our case the "single universe" Hamiltonian. It has eigenfunctions $\psi_n(l)$ with corresponding eigenvalues $e_n = 2n\sqrt{\lambda}$, $n = 1, 2, \ldots$:

$$\psi_n(l) = l\, e^{-\sqrt{\lambda} l}\, p_{n-1}(l), \qquad H_0(l)\psi_n(l) = e_n\psi_n(l), \tag{14.33}$$

where $p_{n-1}(l)$ is a polynomial of order $n-1$. Note that the disk amplitude $W_\lambda^{(0)}(l)$, which is obtained from (14.17), formally corresponds to $n = 0$ in (14.33):

$$W_\lambda^{(0)}(l) = \mathrm{e}^{-\sqrt{\lambda} l}, \quad H_0(l)W_\lambda^{(0)}(l) = 0. \tag{14.34}$$

This last equation can be viewed as a kind of Wheeler–deWitt equation if we view the disk function as the Hartle–Hawking wave function. However, $W_\lambda^{(0)}(l)$ does not belong to the spectrum of $H_0(l)$ since it is not normalizable when one uses the measure (14.32).

We now introduce creation and annihilation operators $a_n^\dagger$ and a_n corresponding to these states, acting on the Fock-vacuum $|0\rangle$ and satisfying $[a_n, a_m^\dagger] = \delta_{n,m}$. We define

$$\Psi(l) = \sum_n a_n\psi_n(l), \quad \Psi^\dagger(l) = \sum_n a_n^\dagger\psi_n^*(l), \tag{14.35}$$

and from the orthonormality of the eigenfunctions with respect to the measure $\mathrm{d}l/l$ we recover (14.31). The "second-quantized" Hamiltonian is

$$\hat{H}_0 = \int_0^\infty \frac{\mathrm{d}l}{l}\, \Psi^\dagger(l)H_0(l)\Psi(l), \tag{14.36}$$

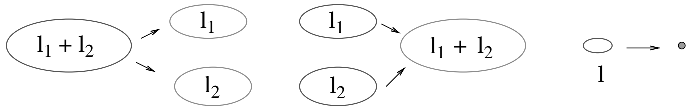


Figure 14.5 Graphical illustration of the various terms in Eq. (14.38).

and the propagator $G_\lambda^{(0)}(l_1, l_2; t)$ is now obtained as

$$G_\lambda^{(0)}(l_1, l_2; t) = \langle 0|\Psi(l_2) \mathrm{e}^{-t\hat{H}_0} \Psi^\dagger(l_1)|0\rangle. \tag{14.37}$$

While this is trivial, the advantage of the formalism is that it automatically takes care of symmetry factors (like in the many-body applications in statistical field theory) both when many spatial universes are at play and when they are joining and splitting. We can follow [17] and define the following Hamiltonian, describing the interaction between spatial universes:

$$\begin{aligned}\hat{H} = \hat{H}_0 - \; & g_s \int \mathrm{d}l_1 \int \mathrm{d}l_2 \Psi^\dagger(l_1)\Psi^\dagger(l_2)\Psi(l_1+l_2) \\ & -\alpha g_s \int \mathrm{d}l_1 \int \mathrm{d}l_2 \Psi^\dagger(l_1+l_2)\Psi(l_2)\Psi(l_1) - \int \frac{\mathrm{d}l}{l}\, \rho(l)\Psi(l),\end{aligned} \tag{14.38}$$

where the different terms of the Hamiltonian are illustrated in Figure 14.5. Here g_s is the coupling constant we have already encountered in Section 14.3, of mass dimension 3. The factor α is just inserted to be able to identify the action of the two g_s-terms in (14.38) when expanding in powers of g_s. We will think of $\alpha = 1$ unless explicitly stated differently. When $\alpha = 1$, $\hat{H}$ is hermitian except for the presence of the tadpole term. It tells us that universes can vanish, but not be created from nothing. The meaning of the two interaction terms is as follows: the first term replaces a universe of length $l_1 + l_2$ with two universes of length l_1 and l_2. This is one of the processes shown in Figure 14.5. The second term represents the opposite process, where two spatial universes merge into one, i.e. the time-reversed picture. The coupling constant g_s clearly appears as a kind of string coupling constant: one factor g_s for splitting spatial universes, one factor g_s for merging spatial universes, and thus a factor g_s^2 when the space-time topology changes, but there are also factors for branching alone. This is only compatible with a Euclidean SFT-picture if we associate a puncture (and thus a topology change) with the vanishing of a baby universe. As discussed above, this is indeed not unnatural from a Lorentzian point of view. From this point of view the appearance of a tadpole term is more natural in the CDT framework than in the original Euclidean framework [17]. The tadpole term is a formal realization of this puncture "process," where the light-cone becomes degenerate.

In principle, we can now calculate the process where we start out with m spatial universes at time 0 and end with n universes at time t, represented as

$$G_{\lambda,g_s}(l_1,..,l_m;l'_1,..,l'_n;t) = \langle 0|\Psi(l'_1)\ldots\Psi(l'_n)\, e^{-t\hat{H}}\Psi^\dagger(l_1)\ldots\Psi^\dagger(l_m)|0\rangle. \tag{14.39}$$

14.5.1 Dyson–Schwinger equations

The disk amplitude is one of a set of functions for which it is possible to derive Dyson–Schwinger equations (DSE). The disk amplitude is characterized by the fact that at $t = 0$ we have a spatial universe of some length, and at some point it vanishes in the "vacuum." Let us consider the more general situation where a set of spatial universes of some lengths l_i exists at time $t = 0$ and where the universes vanish at later times. Define the generating function:

$$Z(J) = \lim_{t\to\infty} \langle 0|\mathrm{e}^{-t\hat{H}}\ \mathrm{e}^{\int \mathrm{d}l\, J(l)\Psi^\dagger(l)}|0\rangle. \tag{14.40}$$

Notice that if the tadpole term had not been present in $\hat{H}$, $Z(J)$ would trivially be equal to 1. We have

$$\lim_{t\to\infty} \langle 0|\mathrm{e}^{-t\hat{H}}\ \Psi^\dagger(l_1)\cdots\Psi^\dagger(l_n)|0\rangle = \left.\frac{\delta^n Z(J)}{\delta J(l_1)\cdots\delta J(l_n)}\right|_{J=0}. \tag{14.41}$$

$Z(J)$ is the generating functional for universes that disappear in the vacuum. We now have

$$0 = \lim_{t\to\infty}\left[\frac{\partial}{\partial t}\ \langle 0|\mathrm{e}^{-t\hat{H}}\ \mathrm{e}^{\int \mathrm{d}l\, J(l)\Psi^\dagger(l)}|0\rangle = -\langle 0|\mathrm{e}^{-t\hat{H}}\ \hat{H}\ \mathrm{e}^{\int \mathrm{d}l\, J(l)\Psi^\dagger(l)}|0\rangle\right]. \tag{14.42}$$

Commuting the $\Psi(l)$'s in $\hat{H}$ past the source term effectively replaces these operators by $lJ(l)$, after which they can be moved to the left of any $\Psi^\dagger(l)$ and outside $\langle 0|$. After that the remaining $\Psi^\dagger(l)$'s in $\hat{H}$ can be replaced by $\delta/\delta J(l)$ and also moved outside $\langle 0|$, leaving us with an integro-differential operator acting on $Z(J)$:

$$0 = \int_0^\infty \mathrm{d}l\, J(l)\, O\left(l, J, \frac{\delta}{\delta J}\right) Z(J), \tag{14.43}$$

where

$$O\left(l, J, \frac{\delta}{\delta J}\right) = H_0(l)\frac{\delta}{\delta J(l)} - \delta(l) - g_s l\int_0^l \mathrm{d}l' \frac{\delta^2}{\delta J(l')\delta J(l-l')}$$
$$-\alpha g_s l \int_0^\infty \mathrm{d}l' l' J(l')\frac{\delta}{\delta J(l+l')}. \tag{14.44}$$

$Z(J)$ is a generating functional which also includes totally disconnected universes which never "interact" with each other. The generating functional for connected universes is obtained in the standard way from field theory by taking the logarithm of $Z(J)$. Thus we write:

$$F(J) = \log Z(J), \tag{14.45}$$

and we have

$$\lim_{t\to\infty} \langle 0|\mathrm{e}^{-t\hat{H}}\Psi^\dagger(l_1)\cdots\Psi^\dagger(l_n)|0\rangle_{con} = \left.\frac{\delta^n F(J)}{\delta J(l_1)\cdots\delta J(l_n)}\right|_{J=0}, \tag{14.46}$$

and we can readily transfer the DSE (14.43)–(14.44) into an equation for the connected functional $F(J)$:

$$\begin{aligned} 0 = \int_0^\infty \mathrm{d}l\, J(l) \Bigg\{ & H_0(l)\frac{\delta F(J)}{\delta J(l)} - \delta(l) - g_s l \int_0^l \mathrm{d}l' \,\frac{\delta^2 F(J)}{\delta J(l')\delta J(l-l')} \\ & - g_s l \int_0^l \mathrm{d}l' \frac{\delta F(J)}{\delta J(l')}\frac{\delta F(J)}{\delta J(l-l')} - \alpha g_s l \int_0^\infty \mathrm{d}l' l' J(l') \frac{\delta F(J)}{\delta J(l+l')} \Bigg\}. \end{aligned} \tag{14.47}$$

From Eq. (14.47) one obtains the DSE by differentiating (14.47) after $J(l)$ a number of times and then taking $J(l) = 0$.

14.5.2 Application of the DSE

Let us introduce the notation

$$w(l_1, \ldots, l_n) \equiv \left.\frac{\delta^n F(J)}{\delta J(l_1)\cdots\delta J(l_n)}\right|_{J=0} \tag{14.48}$$

as well as the Laplace transform $\widetilde{w}(x_1, \ldots, x_n)$. Let us differentiate Eq. (14.47) after $J(l)$ one and two times, then take $J(l) = 0$ and Laplace transform the obtained equations. We obtain the following equations (where $H_0(x) f(x) = \partial_x[(x^2 - \lambda) f(x)]$):

$$0 = H_0(x)\widetilde{w}(x) - 1 + g_s\partial_x\Big(\widetilde{w}(x,x) + \widetilde{w}(x)\widetilde{w}(x)\Big), \tag{14.49}$$

$$\begin{aligned} 0 = {} & (H_0(x) + H_0(y))\widetilde{w}(x,y) + g_s\partial_x\widetilde{w}(x,x,y) + g_s\partial_y\widetilde{w}(x,y,y) \\ & + 2g_s\left(\partial_x[\widetilde{w}(x)\widetilde{w}(x,y)] + \partial_y[\widetilde{w}(y)\widetilde{w}(x,y)]\right) + 2\alpha g_s\partial_x\partial_y\Big(\frac{\widetilde{w}(x) - \widetilde{w}(y)}{x-y}\Big). \end{aligned} \tag{14.50}$$

The structure of the DSE for an increasing number of arguments is hopefully clear (see [16] for details).

We can solve the DSE iteratively. For this purpose let us introduce the expansion of $\widetilde{w}(x_1, \ldots, x_n)$ in terms of the coupling constants g_s and α:

$$\widetilde{w}(x_1, \ldots, x_n) = \sum_{k=n-1}^{\infty} \alpha^k \sum_{m=k-1}^{\infty} g_s^m \, \widetilde{w}(x_1, \ldots, x_n; m, k). \tag{14.51}$$

The amplitude $\widetilde{w}(x_1, \ldots, x_n)$ starts with the power $(\alpha g_s)^{n-1}$ since we have to perform n mergings during the time evolution in order to create a connected geometry if we begin with n separated spatial loops. Thus one can find the lowest-order contribution to $\widetilde{w}(x_1)$ from (14.49), use that to find the lowest-order contribution to $\widetilde{w}(x_1, x_2)$ from (14.50), etc. Returning to Eq. (14.49) we can use the lowest-order expression for $\widetilde{w}(x_1, x_2)$ to find the next-order correction to $\widetilde{w}(x_1)$, etc.

As mentioned above, the amplitude $\widetilde{w}(x_1, \ldots, x_n)$ starts with the power $(\alpha g_s)^{n-1}$ coming from merging the n disconnected spatial universes. The rest of the powers of αg_s will result in a topology change of the resulting, connected worldsheet. From a Euclidean point of view it is thus more appropriate to reorganize the series as follows:

$$\widetilde{w}(x_1, \ldots, x_n) = (\alpha g_s)^{n-1} \sum_{h=0}^{\infty} (\alpha g_s^2)^h \, \widetilde{w}_h(x_1, \ldots, x_n), \tag{14.52}$$

$$\widetilde{w}_h(x_1, \ldots, x_n) = \sum_{j=0}^{\infty} g_s^j \, \widetilde{w}(x_1, \ldots, x_n; n-1+2h+j, n-1+h), \tag{14.53}$$

and aim for a topological expansion in αg_s^2, at each order solving for all possible baby-universe creations which at some point will vanish into the vacuum. Thus $\widetilde{w}_h(x_1, \ldots, x_n)$ will be a function of g_s, although we do not write it explicitly. The DSE allow us to obtain the topological expansion iteratively, much the same way we already did as a power expansion in g_s.

14.6 The matrix model, once again

Let us consider our $N \times N$ hermitian matrix with the cubic potential (14.28) and define the observable

$$\widetilde{W}(x_1, \ldots, x_n)_d = N^{n-2} \left\langle \mathrm{Tr} \left(\frac{1}{x_1 - M} \right) \cdots \left(\mathrm{tr} \, \frac{1}{x_1 - M} \right) \right\rangle, \tag{14.54}$$

where the subscript d refers to the fact that the correlator will contain disconnected parts. We denote the connected part of the correlator by $\widetilde{W}(x_1, \ldots, x_n)$. It is standard matrix model technology to find the matrix model DSE for $\widetilde{W}(x_1, \ldots, x_n)$. We refer to [18–21] for details. *One obtains precisely the same set of coupled equations as (14.49)–(14.50) if we identify*:

$$\alpha = \frac{1}{N^2}, \tag{14.55}$$

and the discussion surrounding the expansion (14.52) is nothing but the standard discussion of the large-N expansion of the multi-loop correlators (see for instance [21] or the more recent papers [22–24]). Thus we conclude that there is a perturbative agreement between the matrix model (14.28) and the CDT SFT in the sense that perturbatively:

$$\widetilde{W}(x_1,\ldots,x_n) = \widetilde{w}(x_1,\ldots,x_n). \tag{14.56}$$

In practice the SFT *is* only defined perturbatively, although in principle we have available the string field Hamiltonian. However, we can now use the matrix model to extract non-perturbative information. The identification of the matrix model and the CDT SFT DSEs were based on (14.55), but in the SFT we are interested in $\alpha = 1$, i.e. formally in $N = 1$, in which case the matrix integrals reduce to ordinary integrals. This means that we will consider the entire sum over topologies "in one go":

$$Z(g,\lambda) = \int \mathrm{d}m \, \exp\left[-\frac{1}{g_s}\left(\lambda m - \frac{1}{3}\, m^3\right)\right], \tag{14.57}$$

while the observables (14.54) can be written as

$$\widetilde{W}_d(x_1,\ldots,x_n) = \frac{1}{Z(g_s,\lambda)} \int \mathrm{d}m \, \frac{\exp\left[-\frac{1}{g_s}\left(\lambda m - \frac{1}{3}\, m^3\right)\right]}{(x_1 - m)\cdots(x_n - m)}. \tag{14.58}$$

These integrals should be understood as formal power series in the dimensionless variable κ defined by Eq. (14.18). Any choice of an integration contour which makes the integral well defined and reproduces the formal power series is a potential non-perturbative definition of these observables. However, different contours might produce different non-perturbative contributions (i.e. which cannot be expanded in powers of t), and there may even be non-perturbative contributions which are not captured by any choice of integration contour. As usual in such situations, additional physics input is needed to fix these contributions.

To illustrate the point, let us start by evaluating the partition function given in (14.57). We have to decide on an integration path in the complex plane in order to define the integral. One possibility is to take a path along the negative axis and then along either the positive or the negative imaginary axis. The corresponding integrals are

$$Z(g_s,\lambda) = \sqrt{\lambda}\,\kappa^{1/3} F_{\pm}(\kappa^{-2/3}), \quad F_{\pm}(\kappa^{-2/3}) = 2\pi\, e^{\pm i\pi/6}\mathrm{Ai}(\kappa^{-2/3}\mathrm{e}^{\pm 2\pi i/3}), \tag{14.59}$$

where Ai denotes the Airy function. Both $F_\pm$ have the same asymptotic expansion in κ, with positive coefficients. Had we chosen the integration path entirely along the imaginary axis we would have obtained ($2\pi i$ times) $\mathrm{Ai}(\kappa^{-2/3})$, but this has an asymptotic expansion in κ with coefficients of oscillating sign, which is at odds with its interpretation as a probability amplitude. In the notation of [25] we have

$$F_\pm(z) = \pi\Big(\mathrm{Bi}(z) \pm i\,\mathrm{Ai}(z)\Big), \tag{14.60}$$

from which one deduces immediately that the functions $F_\pm(\kappa^{-2/3})$ are not real. However, since $\mathrm{Bi}(\kappa^{-2/3})$ grows like $\mathrm{e}^{\frac{2}{3\kappa}}$ for small κ while $\mathrm{Ai}(\kappa^{-2/3})$ falls off like $\mathrm{e}^{-\frac{2}{3\kappa}}$, their imaginary parts are exponentially small in $1/\kappa$ compared to the real part, and therefore do not contribute to the asymptotic expansion in κ. An obvious way to *define* a partition function which is real and shares the same asymptotic expansion is by symmetrization:

$$\frac{1}{2}(F_+ + F_-) \equiv \pi\,\mathrm{Bi}. \tag{14.61}$$

The situation parallels the one encountered in the double-scaling limit of the "old" matrix model [26], and discussed in detail in [27], but is less complicated. We will return to a discussion of this in Section 14.8.

Presently, let us collectively denote by $F(z)$ any of the functions $F_\pm(z)$ or $\pi\,\mathrm{Bi}(z)$, leading to the tentative identification

$$Z(g_s,\lambda) = \sqrt{\lambda}\,\kappa^{1/3}\,F\Big(\kappa^{-2/3}\Big), \quad F''(z) = zF(z), \tag{14.62}$$

where we have included the differential equation satisfied by the Airy functions for later reference. Remarkably, this partition function was also found in [28], where a double-scaling limit of so-called branched polymers was studied. It reflects that a significant part of the dynamics associated with the branching, as reflected in Figures 14.3 and 14.4, is indeed captured by a branched polymer model. It should be no surprise that this is possible. Branched polymers play an important role in non-critical string theory [29, 30] and even in higher-dimensional Euclidean quantum gravity [31]. However, the 2D CDT has a much richer set of observables than those encountered in the theory of branched polymers, for instance the observables $\widetilde{W}_d(x_1,\ldots,x_n)$, the calculation of which we now turn to.

Let us introduce the dimensionless variables

$$x = \widetilde{x}\sqrt{\lambda}, \ \ m = g_s^{1/3}\beta, \quad \widetilde{W}_d(x_1,\ldots,x_n) = \lambda^{-n/2}\widetilde{w}_d(\widetilde{x}_1,\ldots,\widetilde{x}_n). \tag{14.63}$$

Assuming $\widetilde{x}_k > 0$, we can write

$$\frac{1}{\widetilde{x}-\kappa^{1/3}\beta} = \int_0^\infty \mathrm{d}\alpha\, \exp\Big[-\Big(\widetilde{x}-\kappa^{1/3}\beta\Big)\,\alpha\Big]. \tag{14.64}$$

We can use this identity to re-express the pole terms in Eq. (14.58) to obtain the integral representation

$$\widetilde{w}_d(\widetilde{x}_1,\ldots,\widetilde{x}_n) = \int_0^\infty \prod_{i=1}^n \mathrm{d}\alpha_i \; \mathrm{e}^{-(\widetilde{x}_1\alpha_1+\cdots+\widetilde{x}_n\alpha_n)} \; \frac{F\left(\kappa^{-\frac{2}{3}} - \kappa^{\frac{1}{3}}\sum_{i=1}^n \alpha_i\right)}{F\left(\kappa^{-\frac{2}{3}}\right)} \tag{14.65}$$

for the amplitude with dimensionless arguments. By an inverse Laplace transformation we thus obtain:

$$W_d(l_1,\ldots,l_n) = \frac{F(\kappa^{-2/3} - \kappa^{1/3}\sqrt{\lambda}\,(l_1+\cdots+l_n))}{F(\kappa^{-2/3})}. \tag{14.66}$$

For the special case $n = 1$ we find

$$W(l) = \frac{F(\kappa^{-2/3} - \kappa^{1/3}\sqrt{\lambda}\,l)}{F(\kappa^{-2/3})} \tag{14.67}$$

for the disk amplitude, together with the remarkable relation

$$W_d(l_1,\ldots,l_n) = W(l_1+\cdots+l_n). \tag{14.68}$$

By Laplace transformation this formula implies the relation

$$\widetilde{W}_d(x_1,\ldots,x_n) = \sum_{i=1}^n \frac{\widetilde{W}(x_i)}{\prod_{j\neq i}^n (x_j - x_i)}. \tag{14.69}$$

From $\widetilde{W}_d(x_1,\ldots,x_n)$ we can construct the connected multiloop functions $\widetilde{W}(x_1,\ldots,x_n)$ using standard field theory. Let us remark that the asymptotic expansion in κ of $\widetilde{W}(x_1,\ldots,x_n)$ of course agrees with that obtained by recursively solving the CDT Dyson–Schwinger equations.

14.7 Stochastic quantization

It is a most remarkable fact that the above-mentioned results can all be understood as a result of stochastic quantization of *space*. In this picture time becomes the *stochastic time* related to the branching of space into baby universes and the original CDT model described in Section 14.2 becomes the classical limit where no stochastic processes are present [32].

Recall the Langevin stochastic differential equation for a single variable x (see, for example, [33, 34]):

$$\dot{x}^{(\nu)}(t) = -f\left(x^{(\nu)}(t)\right) + \sqrt{\Omega}\,\nu(t), \tag{14.70}$$

where the dot denotes differentiation with respect to stochastic time t, $\nu(t)$ is a Gaussian white-noise term of unit width, and $f(x)$ a dissipative drift force:

$$f(x) = \frac{\partial S(x)}{\partial x}. \tag{14.71}$$

The noise term creates a probability distribution of $x(t)$, reflecting the assumed stochastic nature of the noise term, with an associated probability distribution

$$P(x, x_0; t) = \Big\langle \delta(x - x^{(\nu)}(t; x_0)) \Big\rangle_\nu, \tag{14.72}$$

where the expectation value refers to an average over the Gaussian noise. $P(x, x_0; t)$ satisfies the Fokker–Planck equation

$$\frac{\partial P(x, x_0; t)}{\partial t} = \frac{\partial}{\partial x}\left(\frac{1}{2}\Omega \frac{\partial P(x, x_0; t)}{\partial x} + f(x) P(x, x_0; t)\right). \tag{14.73}$$

This is an imaginary-time Schrödinger equation, with $\sqrt{\Omega}$ playing a role similar to $\hbar$. It enables us to write P as a propagator for a Hamiltonian operator $\hat{H}$,

$$P(x, x_0; t) = \langle x | \mathrm{e}^{-t\hat{H}} | x_0 \rangle, \quad \hat{H} = \frac{1}{2}\Omega \hat{p}^2 + i\, \hat{p}\, f(\hat{x}), \tag{14.74}$$

with initial condition $x(t = 0) = x_0$, and $\hat{p} = -i\partial_x$. It follows that by defining

$$\widetilde{G}(x_0, x; t) \equiv \frac{\partial}{\partial x_0} P(x, x_0; t), \tag{14.75}$$

the function $\widetilde{G}(x_0, x; t)$ satisfies the differential equation

$$\frac{\partial \widetilde{G}(x_0, x; t)}{\partial t} = \frac{\partial}{\partial x_0}\left(\frac{1}{2}\Omega \frac{\partial \widetilde{G}(x_0, x; t)}{\partial x_0} - f(x_0)\, \widetilde{G}(x_0, x; t)\right). \tag{14.76}$$

Omitting the noise term corresponds to taking the limit $\Omega \to 0$. One can then drop the functional average over the noise in (14.72) to obtain

$$P_{cl}(x, x_0; t) = \delta(x - x(t, x_0)), \quad \widetilde{G}_{cl}(x_0, x; t) = \frac{\partial}{\partial x_0}\delta(x - x(t, x_0)). \tag{14.77}$$

It is readily verified that these functions satisfy Eqs (14.73) and (14.76) with $\Omega = 0$. Thus we have for $S(x) = -\lambda x + x^3/3$:

$$\frac{\partial \widetilde{G}_{cl}(x_0, x; t)}{\partial t} = \frac{\partial}{\partial x_0}\Big((\lambda - x_0^2)\, \widetilde{G}_{cl}(x_0, x; t)\Big). \tag{14.78}$$

Comparing now Eqs (14.7) and (14.78), we see that we can formally reinterpret $\widetilde{G}^{(0)}_\lambda(x_0, x; t)$ – an amplitude obtained by non-perturbatively quantizing Lorentzian pure gravity in two dimensions – as the "*classical* probability" $\widetilde{G}_{cl}(x_0, x; t)$ corresponding to the action $S(x) = -\lambda x + x^3/3$ of a zero-dimensional system in the context of stochastic quantization. Only the boundary condition is different, since in the case of CDT x is not an ordinary real variable, but the cosmological constant. The correct boundary conditions are thus the ones stated in Eqs (14.8)–(14.9).

Stochastic quantization of the system amounts to replacing

$$\widetilde{G}_{\lambda}^{(0)}(x_0, x; t) \to \widetilde{G}(x_0, x; t), \tag{14.79}$$

where $\widetilde{G}(x_0, x; t)$ satisfies the differential equation corresponding to Eq. (14.76), namely

$$\frac{\partial \widetilde{G}(x_0, x; t)}{\partial t} = \frac{\partial}{\partial x_0}\left(g_s \frac{\partial}{\partial x_0} + \lambda - x_0^2\right)\widetilde{G}(x_0, x; t). \tag{14.80}$$

We have introduced the parameter $g_s := \Omega/2$, which will allow us to reproduce the matrix model and SFT results reported above.

A neat geometric interpretation of how stochastic quantization can capture topologically non-trivial amplitudes has been given in [35]. Applied to the present case, we can view the propagation in stochastic time t for a given noise term $\nu(t)$ as classical in the sense that solving the Langevin equation (14.70) for $x^{(\nu)}(t)$ iteratively gives precisely the tree diagrams with one external leg corresponding to the action $S(x)$ (and including the derivative term $\dot{x}^{(\nu)}(t)$), with the noise term acting as a source term. Performing the functional integration over the Gaussian noise term corresponds to integrating out the sources and creating loops, or, if we have several independent trees, to merging these trees and creating diagrams with several external legs. If the dynamics of the quantum states of the spatial universe takes place via the strictly causal CDT-propagator $\hat{G}_0 = \mathrm{e}^{-t\hat{H}_0}$, a single spatial universe of length l cannot split into two spatial universes. Similarly, no two spatial universes are allowed to merge as a function of stochastic time. However, introducing the noise term *and* subsequently performing a functional integration over it makes these processes possible. This explains how the stochastic quantization can automatically generate the amplitudes which are introduced by hand in a string field theory, be it of Euclidean character as described in [35], or within the framework of CDT.

What is new in the CDT string field theory considered here is that we can use the corresponding stochastic field theory to solve the model, since we arrive at closed equations valid to all orders in the genus expansion. Let us translate Eq. (14.80) to l-space:

$$\frac{\partial G(l_0, l; t)}{\partial t} = -H(l_0)\, G(l_0, l; t), \tag{14.81}$$

where the *extended* Hamiltonian

$$H(l) = -l\frac{\partial^2}{\partial l^2} + \lambda l - g_s l^2 = H_0(l) - g_s l^2 \tag{14.82}$$

now has an extra potential term coming from the inclusion of branching points compared to the Hamiltonian $H_0(l)$ defined in (14.14). It is truly remarkable that all branching and joining is contained in this simple extra term. Formally $H(l)$ is a well-defined hermitian operator with respect to the measure (14.32) (we will discuss some subtleties in the next section).

We can now write down the generalization of the Wheeler–deWitt equation (14.32) for the disk amplitude:

$$H(l)W(l) = 0. \tag{14.83}$$

Contrary to $W_\lambda^{(0)}(l)$ appearing in (14.32), $W(l)$ contains all branchings and all topology changes, and the solution is precisely (14.67)! This justifies the choice $g_s = \Omega/2$ mentioned above. Recall that $E = 0$ does not belong to the spectrum of $H_0(l)$ since $W_0(l)$ is not integrable at zero with respect to the measure (14.32). Exactly the same is true for the extended Hamiltonian $H(l)$ and the corresponding Hartle–Hawking amplitude $W(l)$.

We also have as a generalization of (14.13) that

$$G(l_0, l; t) = \langle l|\mathrm{e}^{-tH(l)}|l_0\rangle \tag{14.84}$$

describes the non-perturbative propagation of a spatial loop of length l_0 to a spatial loop of length l in proper (or stochastic) time t, now including the summation over all genera.

14.8 The extended Hamiltonian

In order to analyze the spectrum of $H(l)$, it is convenient to put the differential operator into standard form. After a change of variables

$$l = \frac{1}{2}z^2, \qquad \psi(l) = \sqrt{z}\phi(z), \tag{14.85}$$

the eigenvalue equation becomes

$$H(z)\phi(z) = E\phi(z), \quad H(z) = -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}z^2} + \frac{1}{2}\lambda z^2 + \frac{3}{8z^2} - \frac{g_s}{4}z^4. \tag{14.86}$$

This shows that the potential is unbounded from below, but such that the eigenvalue spectrum is still discrete: whenever the potential is unbounded below with fall-off faster than $-z^2$, the spectrum is discrete, reflecting the fact that the classical escape time to infinity is finite (see [36] for a detailed discussion relevant to the present situation). For small g_s, there is a large barrier of height $\lambda^2/(2g_s)$ separating the unbounded region for $l > \lambda/g_s$ from the region $0 \le l \le \lambda/(2g_s)$ where the potential grows. This situation is perfectly suited to applying a standard WKB analysis. For energies less than $\lambda^2/(2g_s)$, the eigenfunctions (14.33) of $H_0(l)$ will be good approximations to those of $\hat{H}(l)$. However, when $l > \lambda/g_s$ the exponential fall-off of $\psi_n^{(0)}(l)$ will be replaced by an oscillatory behavior, with the wave function falling off only like $1/l^{1/4}$. The corresponding $\psi_n(l)$ is still square-integrable since we have to use the measure (14.32). For energies larger than $\lambda^2/(2g_s)$, the solutions will be entirely oscillatory, but still square-integrable.

Thus a somewhat drastic change has occurred in the quantum behavior of the one-dimensional universe as a consequence of allowing topology changes. In the original, strictly causal quantum gravity model an eigenstate $\psi_n^{(0)}(l)$ of the spatial universe had an average size of order $1/\sqrt{\lambda}$. However, allowing for branching and topology change, the average size of the universe is now infinite!

As discussed in [36], Hamiltonians with unbounded potentials like (14.86) have a one-parameter family of self-adjoint extensions and we still have to choose one of those such that the spectrum of $H(l)$ can be determined unambiguously. One way of doing this is to appeal again to stochastic quantization, following the strategy used by Greensite and Halpern [37], which was applied to the double-scaling limit of matrix models in [36, 38, 39]. The Hamiltonian (14.74) corresponding to the Fokker–Planck equation (14.80), namely

$$H(x)\psi(x) = -g_s \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}S(x)}{\mathrm{d}x}\psi(x)\right), \quad S(x) = \left(\frac{x^3}{3} - \lambda x\right), \tag{14.87}$$

is not hermitian if we view x as an ordinary real variable and wave functions $\psi(x)$ as endowed with the standard scalar product on the real line. However, by a similarity transformation one can transform $H(x)$ to a new operator

$$\widetilde{H}(x) = \mathrm{e}^{-S(x)/2g_s} H(x)\, \mathrm{e}^{S(x)/2g_s}; \quad \widetilde{\psi}(x) = \mathrm{e}^{-S(x)/2g_s}\psi(x), \tag{14.88}$$

which *is* hermitian on $L^2(R, \mathrm{d}x)$. We have

$$\widetilde{H}(x) = -g_s \frac{\mathrm{d}^2}{\mathrm{d}x^2} + \left(\frac{1}{4g_s}\left(\frac{\mathrm{d}S(x)}{\mathrm{d}x}\right)^2 + \frac{1}{2}\frac{\mathrm{d}^2S(x)}{\mathrm{d}x^2}\right), \tag{14.89}$$

which after substitution of the explicit form of the action becomes

$$\widetilde{H}(x) = -g_s \frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x), \quad V(x) = \frac{1}{4g_s}(\lambda - x^2)^2 + x. \tag{14.90}$$

The fact that one can write

$$\widetilde{H}(x) = R^\dagger R, \quad R = -\sqrt{g_s}\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2\sqrt{g_s}}\frac{\mathrm{d}S(x)}{\mathrm{d}x} \tag{14.91}$$

implies that the spectrum of $\widetilde{H}(x)$ is positive, discrete and unambiguous. We conclude that the formalism of stochastic quantization has provided us with a non-perturbative definition of the CDT string field theory.

Acknowledgments

JA, RL, WW and SZ acknowledge support by ENRAGE (European Network on Random Geometry), a Marie Curie Research Training Network, contract MRTN-CT-2004-005616, and RL acknowledges support by the Netherlands Organisation for Scientific Research (NWO) under their VICI program. SZ thanks the Department of Statistics at Sao Paulo University for kind hospitality and acknowledges financial support of the ISAC program, Erasmus Mundus.

References

[1] S. Weinberg: Ultraviolet divergences in quantum theories of gravitation, in *General Relativity: Einstein Centenary Survey*, eds. S. W. Hawking and W. Israel, Cambridge University Press, Cambridge, UK (1979) 790–831.

[2] A. Codello, R. Percacci, and C. Rahmede, Investigating the ultraviolet properties of gravity with a Wilsonian renormalization group equation, *Annals Phys.* **324** (2009) 414 [arXiv:0805.2909 [hep-th]]. M. Reuter and F. Saueressig: Functional renormalization group equations, asymptotic safety, and quantum Einstein gravity [0708.1317, hep-th]. M. Niedermaier and M. Reuter: The asymptotic safety scenario in quantum gravity, *Living Rev. Rel.* **9** (2006) 5. H. W. Hamber and R. M. Williams: Nonlocal effective gravitational field equations and the running of Newton's G, *Phys. Rev. D* **72** (2005) 044026 [hep-th/0507017]. D. F. Litim: Fixed points of quantum gravity, *Phys. Rev. Lett.* **92** (2004) 201301 [hep-th/0312114]. H. Kawai, Y. Kitazawa, and M. Ninomiya: Renormalizability of quantum gravity near two dimensions, *Nucl. Phys. B* **467** (1996) 313–31 [hep-th/9511217].

[3] J. Ambjørn, J. Jurkiewicz, and R. Loll: Dynamically triangulating Lorentzian quantum gravity, *Nucl. Phys. B* **610** (2001) 347–82 [hep-th/0105267].

[4] J. Ambjørn, J. Jurkiewicz, and R. Loll: Reconstructing the universe, *Phys. Rev. D* **72** (2005) 064014 [hep-th/0505154].

[5] J. Ambjørn, A. Görlich, J. Jurkiewicz, and R. Loll, The nonperturbative quantum de Sitter universe, *Phys. Rev. D* **78** (2008) 063544 [arXiv:0807.4481 [hep-th]]. J. Ambjorn, A. Görlich, J. Jurkiewicz and R. Loll: Planckian birth of the quantum de Sitter universe, *Phys. Rev. Lett.* **100** (2008) 091304 [0712.2485, hep-th].

[6] J. Ambjørn, J. Jurkiewicz, and R. Loll: Emergence of a 4D world from causal quantum gravity, *Phys. Rev. Lett.* **93** (2004) 131301 [hep-th/0404156].

[7] J. Ambjørn, J. Jurkiewicz, and R. Loll: The universe from scratch, *Contemp. Phys.* **47** (2006) 103–17 [hep-th/0509010]. R. Loll: The emergence of spacetime, or, quantum gravity on your desktop, *Class. Quant. Grav.* **25** (2008) 114006 [0711.0273, gr-qc].

[8] J. Ambjørn, J. Jurkiewicz, and R. Loll: Semiclassical universe from first principles, *Phys. Lett. B* **607** (2005) 205–13 [hep-th/0411152]. Spectral dimension of the universe, *Phys. Rev. Lett.* **95** (2005) 171301 [hep-th/0505113].

[9] O. Lauscher and M. Reuter: Fractal spacetime structure in asymptotically safe gravity, *JHEP* **0510** (2005) 050 [arXiv:hep-th/0508202].

[10] P. Horava: Spectral dimension of the universe in quantum gravity at a Lifshitz point, *Phys. Rev. Lett.* **102** (2009) 161301 [arXiv:0902.3657 [hep-th]].

[11] T. Regge: General relativity without coordinates, *Nuovo Cim.* **19** (1961) 558.

[12] J. Ambjørn and R. Loll: Non-perturbative Lorentzian quantum gravity, causality and topology change, *Nucl. Phys. B* **536** (1998) 407–34 [hep-th/9805108].

[13] C. Teitelboim: Causality versus gauge invariance in quantum gravity and supergravity, *Phys. Rev. Lett.* **50** (1983) 705–8. The proper time gauge in quantum theory of gravitation, *Phys. Rev. D* **28** (1983) 297–309.

[14] J. Ambjørn, R. Loll, W. Westra, and S. Zohren: Putting a cap on causality violations in CDT, *JHEP* **0712** (2007) 017 [0709.2784, gr-qc].

[15] J. Ambjørn, R. Loll, Y. Watabiki, W. Westra, and S. Zohren, A matrix model for 2D quantum gravity defined by causal dynamical triangulations, *Phys. Lett. B* **665** (2008) 252–56 [0804.0252, hep-th]. A new continuum limit of matrix models, *Phys. Lett. B* **670** (2008) 224 [arXiv:0810.2408 [hep-th]]. A causal alternative for $c = 0$ strings, *Acta Phys. Polon. B* **39** (2008) 3355 [arXiv:0810.2503 [hep-th]].

[16] J. Ambjørn, R. Loll, Y. Watabiki, W. Westra, and S. Zohren: A string field theory based on causal dynamical triangulations, *JHEP* **0805** (2008) 032 [0802.0719, hep-th].

[17] H. Kawai, N. Kawamoto, T. Mogami, and Y. Watabiki: Transfer matrix formalism for two-dimensional quantum gravity and fractal structures of space-time, *Phys. Lett. B* **306** (1993) 19–26 [hep-th/9302133]. N. Ishibashi and H. Kawai: String field theory of noncritical strings, *Phys. Lett. B* **314** (1993) 190 [arXiv:hep-th/9307045]. String field theory of $c \leq 1$ noncritical strings, *Phys. Lett. B* **322** (1994) 67 [arXiv:hep-th/9312047]. A background independent formulation of noncritical string theory, *Phys. Lett. B* **352** (1995) 75 [arXiv:hep-th/9503134]. M. Ikehara, N. Ishibashi, H. Kawai, T. Mogami, R. Nakayama, and N. Sasakura: String field theory in the temporal gauge, *Phys. Rev. D* **50** (1994) 7467 [arXiv:hep-th/9406207]. Y. Watabiki: Construction of noncritical string field theory by transfer matrix formalism in dynamical triangulation, *Nucl. Phys. B* **441** (1995) 119–66 [hep-th/9401096]. H. Aoki, H. Kawai, J. Nishimura, and A. Tsuchiya: Operator product expansion in two-dimensional quantum gravity, *Nucl. Phys. B* **474** (1996) 512–28 [hep-th/9511117]. J. Ambjørn and Y. Watabiki: Noncritical string field theory for 2D quantum gravity coupled to (p, q)–conformal fields, *Int. J. Mod. Phys. A* **12** (1997) 4257 [arXiv:hep-th/9604067].

[18] F. David: Loop equations and nonperturbative effects in two-dimensional quantum gravity, *Mod. Phys. Lett. A* **5** (1990) 1019.

[19] J. Ambjørn, J. Jurkiewicz, and Yu. M. Makeenko: Multiloop correlators for two-dimensional quantum gravity, *Phys. Lett. B* **251** (1990) 517–24.

[20] J. Ambjørn and Yu. M. Makeenko: Properties of loop equations for the Hermitean matrix model and for two-dimensional quantum gravity,' *Mod. Phys. Lett. A* **5** (1990) 1753.

[21] J. Ambjørn, L. Chekhov, C. F. Kristjansen, and Yu. Makeenko: Matrix model calculations beyond the spherical limit, *Nucl. Phys. B* **404** (1993) 127 [Erratum-ibid. **449** (1995) 681] [arXiv:hep-th/9302014].

[22] B. Eynard: Topological expansion for the 1-hermitian matrix model correlation functions, *JHEP* **0411** (2004) 031 [arXiv:hep-th/0407261].

[23] L. Chekhov and B. Eynard: Hermitean matrix model free energy: Feynman graph technique for all genera, *JHEP* **0603** (2006) 014 [arXiv:hep-th/0504116].

[24] B. Eynard and N. Orantin: Invariants of algebraic curves and topological expansion, arXiv:math-ph/0702045. Topological expansion and boundary conditions, *JHEP* **0806** (2008) 037 [arXiv:0710.0223 [hep-th]].

[25] M. Abramowitz and I. Stegun (eds): *Pocketbook of Mathematical Functions* (Harri Deutsch, Frankfurt, 1984).

[26] F. David: Nonperturbative effects in matrix models and vacua of two-dimensional gravity, *Phys. Lett. B* **302** (1993) 403 [hep-th/9212106].

[27] M. Mariño: Nonperturbative effects and nonperturbative definitions in matrix models and topological strings, *JHEP* **0812** (2008) 114 [0805.3033 [hep-th]].

[28] J. Jurkiewicz and A. Krzywicki: Branched polymers with loops, *Phys. Lett. B* **392** (1997) 291 [hep-th/9610052].

[29] J. Ambjørn and B. Durhuus: Regularized bosonic strings need extrinsic curvature, *Phys. Lett. B* **188** (1987) 253–57.

[30] J. Ambjørn, S. Jain, and G. Thorleifsson: Baby universes in 2-d quantum gravity, *Phys. Lett. B* **307** (1993) 34–9 [hep-th/9303149].

[31] J. Ambjørn, S. Jain, J. Jurkiewicz, and C. F. Kristjansen: Observing 4-d baby universes in quantum gravity, *Phys. Lett. B* **305** (1993) 208 [hep-th/9303041].

[32] J. Ambjørn, R. Loll, W. Westra, and S. Zohren: Stochastic quantization and the role of time in quantum gravity, *Phys. Lett. B* **680** (2009) 359 [arXiv:0908.4224 [hep-th]].

[33] J. Zinn-Justin: Quantum field theory and critical phenomena, *Int. Ser. Monogr. Phys.* **113** (2002) 1.

[34] M. Chaichian and A. Demichev: *Path Integrals in Physics, Volume II*, Institute of Physics Publishing, Bristol, UK (2001).

[35] M. Ikehara, N. Ishibashi, H. Kawai, T. Mogami, R. Nakayama, and N. Sasakura: A note on string field theory in the temporal gauge, *Prog. Theor. Phys. Suppl.* **118** (1995) 241 [arXiv:hep-th/9409101].

[36] J. Ambjørn and C. F. Kristjansen: Nonperturbative 2-d quantum gravity and Hamiltonians unbounded from below, *Int. J. Mod. Phys. A* **8** (1993) 1259 [arXiv: hep-th/9205073].

[37] J. Greensite and M. B. Halpern: Stabilizing bottomless action theories, *Nucl. Phys. B* **242** (1984) 167.

[38] J. Ambjørn and J. Greensite: Nonperturbative calculation of correlators in 2-D quantum gravity, *Phys. Lett. B* **254** (1991) 66.

[39] J. Ambjørn, J. Greensite, and S. Varsted: A nonperturbative definition of 2-D quantum gravity by the fifth time action, *Phys. Lett. B* **249** (1990) 411.

15

Logic is to the quantum as geometry is to gravity

RAFAEL SORKIN

I will propose that the reality to which the quantum formalism implicitly refers is a kind of generalized history, the word history having here the same meaning as in the phrase sum-over-histories. This proposal confers a certain independence on the concept of event, and it modifies the rules of inference concerning events in order to resolve a contradiction between the idea of reality as a single history and the principle that events of zero measure cannot happen (the Kochen–Specker paradox being a classic expression of this contradiction). The so-called measurement problem is then solved if macroscopic events satisfy classical rules of inference, and this can in principle be decided by a calculation. The resulting conception of reality involves neither multiple worlds nor external observers. It is therefore suitable for quantum gravity in general and causal sets in particular.

15.1 Quantum gravity and quantal reality

Why, in our attempts to unify our theories of gravity and the quantum, has progress been so slow? One reason, no doubt, is that it's simply a very hard problem. Another is that we lack clear guidance from experiments or astronomical observations. But I believe that a third thing holding us back is that we haven't learned how to think clearly about the quantum world in itself, without reference to "observers" and other external agents.

Because of this we don't really know how to think about the Planckian regime where quantum gravity is expected to be most relevant. We don't know how to think about the vacuum on small scales, or about the inside of a black hole, or

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

about the early universe. Nor do we have a way to pose questions about relativistic causality in such situations. This is particularly troubling for the causal set program [1], within which a condition of "Bell causality" has been defined in the classical case, and has led there to a natural family of dynamical laws (those of the CSG or "classical sequential growth" models) [2]. If we possessed an analogous concept of "quantal Bell causality," we could set about deriving a dynamics of quantal sequential growth. But without an observer-free notion of reality, how does one give meaning to superluminal causation or its absence in a causal set?

It's not that individual physicists have no notion of what the quantal world is like, of course. We all employ intuitive pictures in our work, and for example, I imagine that very few people think of rotons in a superfluid in terms of self-adjoint operators. But what we lack is a coherent descriptive framework. We lack, in other words, an answer to the question, "What is a quantal reality?"

My main purpose in this chapter is first to propose an answer to this question (or really a family of possible answers), and second to explain how, on the basis of this answer, the so-called measurement problem can be posed and plausibly solved. My proposal belongs to the histories-based way of thinking about dynamics, which in a quantal context corresponds to path-integral formulations. More specifically, it rests on three or four basic ideas, that of *event*, that of *preclusion*, and that of *anhomomorphic inference* concerning *coevents*, whose meaning I will try to clarify in what follows.

15.2 Histories and events (the kinematic input)

In classical physics, it was easy to say what a possible reality was, although the form of the answer was not static, but changed as our knowledge of nature grew. Electromagnetic theory, for example, conceived reality as a background Minkowski spacetime inhabited by a Faraday field F_{ab} together with a collection of particle worldlines, each with a given charge and mass, while reality for general relativity was a 4-geometry together with possible matter fields, thus a diffeomorphism-equivalence class of Lorentzian metrics and other fields. Of course, we were far from knowing all the details of the actual reality, but we could say exactly what in principle it would have taken to describe reality fully if we did have the details. Thus we could survey all the kinematically possible realities, and then go on to state the dynamical laws (equations of motion or field equations) that further circumscribed these possibilities.

Another example comes from Brownian motion, which in important ways stands closer to quantum mechanics than deterministic classical theories do. Here, if we imagine that nothing exists but one Brownian particle, then a possible reality is just a single worldline (continuous but not differentiable), and the dynamical law is a set of transition probabilities, or more correctly a probability measure on the space of all worldlines.

Such a possible reality – a spacetime, a field, a worldline, etc. – is what is meant by the word "history" in the title of this section; and according to the view I am proposing, such histories furnish the raw material from which reality is constructed. The word history thus denotes the same thing it does when people call the path integral a "sum-over-histories." To avoid confusion with other uses of the word, one might say *proto-history* or perhaps "kinematical" or "bare" history. Given this distinction, one might then refer to quantal reality as a "quantal history." However, unlike in classical physics, we will not (in general) identify quantal reality with a single history, instead we will have certain sorts of "logical combinations" of histories which will be described by *coevents*. In the simplest case a coevent will correspond merely to a set of histories. Although, without further preparation, I cannot yet make precise what quantal reality will be, let me stress at the outset what it will *not* be, namely a wave-function or state-vector. Nor will the Schrödinger equation enter the story at all. Such objects can have a technical role to play, but at no stage will they enter the basic interpretive framework.

As already indicated, the concepts of event and coevent will be fundamental to this framework. In order to define them, we need first to introduce the *history space* Ω whose elements are the individual histories. An *event* is then a subset of Ω. When Ω contains an infinite number of histories, not every subset will be an event because some sort of "measurability" condition will be required, but for present purposes I will always assume that $|\Omega| < \infty$. In that case, one can equate the word "event" to the phrase "subset of Ω."

Notice that this definition of the word "event" parallels its use in everyday speech, where a sentence like "It rained all day yesterday" denotes in effect a large number of more detailed specifications of the weather, all lumped together under the heading "rain." (This usage of "event" also follows the customary terminology in probability theory, where however Ω is often called the "sample space" rather than the history space.) On the other hand, one should not confuse event in this sense with the word "event" used to denote a point of spacetime. The latter may also be regarded as a type of event, perhaps, but that would only be a very special case of what I mean by event herein, and it would be relevant only in connection with quantum gravity per se.

Let us write $\mathfrak{A}$ for the space of all events. Structurally, $\mathfrak{A}$ is a Boolean algebra, meaning that union, intersection, complementation, and symmetric-difference are defined for it. In logical terms these correspond respectively to the connectives *or*, *and*, *not*, and *xor*.

Finally, we need to define *coevent*. We have defined the dual object, an event $E \in \mathfrak{A}$, as a subset of Ω, but we can also think of an event as a *question*, in our previous example the question "Did it rain all day yesterday?" A coevent can then be thought of as something that answers all conceivable questions. More formally, a *coevent* will be a map $\phi : \mathfrak{A} \to \mathbb{Z}_2$, where $\mathbb{Z}_2$ is the two-element set $\{0, 1\}$, the intended meaning being that $\phi(E) = 1$ if and only if the event E actually happens. Thus, 1 represents the answer "yes" (or "true") and 0 the answer "no" (or "false"). I will take the point of view that such a coevent is a full description of reality, quantal or classical. After all, what more could one ask for in the way of description than an answer to every question that one might ask about the world? For now we place no conditions on ϕ other than that it takes events to members of $\mathbb{Z}_2$. (Notice that a coevent is in some sense a "higher-order object." If one thought of events as "predicates" then a coevent would be a "predicate of predicates." A function from a Boolean algebra to $\mathbb{Z}_2$ is sometimes called a "truth valuation," but that terminology would be misleading here. It would suggest too strongly that the event-algebra belongs to some *a priori* formal language or logical scheme, with reality being merely a "model" of that logic. But in the context of physics, it would seem that the events come first and the descriptive language second. Moreover, one of the main points of this chapter philosophically is (as its title indicates) that the rules of logical inference are part of physics and can never be written down fully without some knowledge of the dynamics.)

15.3 Preclusion and the quantal measure (the dynamical input)

I've said that anhomomorphic logic grows out of the path integral, but in order to understand what this means, you must think of the path integral as something more than just a propagator from one wave function ψ to another. What, in fact, does the path integral really compute, if we try to understand it on its own terms?

Let us for a (very brief) moment adopt an "operational" point of view which only cares about the probabilities of instrumental "pointer events." Such events can be idealized as "position measurements" (the positions of the pointers), and it has been known for a long time that the joint probabilities for successive position measurements can be computed directly from the path integral without reference to any wave function, except possibly as a shortcut to specifying initial conditions. The probabilities in question here are those furnished by the standard evolve−collapse−evolve algorithm to be found in any textbook of quantum mechanics. When one abstracts from the mathematical machinery used to compute them, what remains is a probability measure μ on a space of "instrument events," and it is this measure, rather than any wave function, that has direct operational meaning! How one may compute μ directly from the path integral is described in more detail in [3], but for us here the important points are three. First that μ refers not to "measurements" per se, but merely to certain macroscopic happenings, and second that it is natural, when μ is expressed as a path integral, to regard these macroscopic happenings as being *events* in precisely the sense defined above. (The histories in this case would specify the trajectories of the molecules comprising the "pointer" and an event would be, as always, a set of histories.) The crucial observation then is that the path-integral computation of μ makes sense for *any* set of histories – any event – and therefore need not be tied to some undefined notion of "measurement."

This event-function $\mu : \mathfrak{A} \to \mathbb{R}^+$ I will call the *quantal measure*, and I will take the point of view that it is the answer to my question above about what the path integral really computes when we try to understand it on its own terms. From this histories vantage point, the textbook rules for computing probabilities are not fundamental principles, but rather rules of thumb whose practical success for certain macroscopic events needs to be explained on the basis of a deeper understanding of what the quantal measure is telling us about microscopic reality. Or to put the point another way, quantum mechanics formulated via the path integral presents itself to us as a generalized probability theory with μ appearing as a generalized probability measure [4–6]. Our first task then is to interpret this generalized measure. (Much more could be said about the formal properties of μ and their relation to the hierarchy of [7] and to the *decoherence functional* that was first defined within the "decoherent histories" interpretation of quantum mechanics [8]. I will, however, limit myself to these two references and to a reference to an experiment currently testing the "3-slit sum rule" that expresses in great generality the quadratic nature of the Born rule [9].)

One's first thought might be to interpret $\mu(E)$ as an ordinary probability attaching to the event E, but this idea founders at once on the failure of μ to be additive on disjoint events. Since such non-additivity expresses the physical phenomenon of *interference* that lies at the heart of quantum mechanics, the impasse seems to me to be definitive: some other concept than probability in the sense of relative frequency seems to be called for. Moreover, μ fails to be bounded above by 1, whence some events E would have to be "more than certain," were we to take $\mu(E)$ as a probability in the ordinary sense. There is, however, one special case in which normalization and additivity become irrelevant, namely for events E such that $\mu(E) = 0$. In such a case, one could conclude on almost any interpretation that the event E should never happen. (Classically, $\mu(E)$ can never vanish exactly except in trivial cases, but quantally it can, thanks precisely to interference!) Such an event (of μ-measure 0) I will denote as *precluded*, and I propose to interpret μ in terms of the following *preclusion postulate*: If $\mu(E) = 0$ then E does not happen.

With respect to a given coevent ϕ, the "not happening" of E is expressed, as we have seen, by the equation $\phi(E) = 0$, and the preclusion postulate thereby becomes a rule limiting the coevents that are dynamically possible:

$$\mu(E) = 0 \quad \Rightarrow \quad \phi(E) = 0.$$

A coevent that fulfills this condition, I will call *preclusive*.

By isolating in this manner what one might call the purely logical implications of the generalized measure μ, one may hope to bring out those aspects which are peculiarly quantal, as opposed to aspects pertaining to probability more generally. Of course it will be necessary at some stage to recover not only the "logical" but also the properly probabilistic predictions one obtains from the standard quantum apparatus. Whether or not the preclusion rule above will suffice for this is not entirely clear, but if it does, one will have cleared up some of the confusion surrounding even the classical probability concept. If, on the other hand, one needed something more than the strict preclusion rule, one could simply extend it to embrace the case of "approximate preclusion," where $\mu(E)$ is not exactly zero but still small enough to be treated as if it vanished. In this way, the difficulties of classical probability would not have grown any better, but (hopefully) they would not have grown any worse either [10]. By basing the probability concept on approximate preclusion, one would in effect be adopting the interpretation sometimes known as Cournot's principle, according to which the assertion that an event of sufficiently small measure will not happen exhausts the scientific meaning of probability. (See [11] for a concise statement of this idea.)[1] Cournot's principle is not free of problems, of course, but neither is any other account of probability, as far as I know. In any case, it seems prudent to leave probability aside at first, and concentrate on the purely logical questions raised by the preclusion principle. Considering that the latter seem to require a radical revision of some basic logical presuppositions, questions of probability might appear in a very different light, once a more adequate picture of quantal reality is in place.

To summarize the burden of this section then, the idea is that the whole dynamical content of the quantal formalism reduces to the preclusion rule stated above (possibly supplemented by its generalization to the case of approximate preclusion).

15.4 The 3-slit paradox and its cognates

Viewed through the lens of the path integral, quantum theory appears as a generalized theory of stochastic processes characterized by the quantal measure μ, and this makes feasible a "histories-based" way of thinking about the dynamics that seems

[1] Predictions about frequencies follow when one construes multiple repetitions of some experiment as a single, combined experiment grouping all the repetitions together into a single sample space. The event that the overall frequencies come out wrong will then possess a tiny measure.

more suited to the needs of quantum gravity than alternative accounts inspired by either the S-matrix or the Schrödinger equation. For such an approach to succeed, however, one needs to free the path integral from its *conceptual* dependence on objects like the wave function. That is, one needs a *free-standing* histories-based formulation of quantum theory. *A priori*, such a formulation need not base itself on the path integral, but as things stand, no other alternative has so far offered itself. In practice then we can (for now at least) vindicate the histories-based viewpoint (also called the "spacetime" viewpoint) only by clarifying the physical meaning of the quantal measure μ.

I have proposed above that the dynamical implications of μ are mediated by what it tells us about the precluded events, the sets of histories of zero measure in Ω. Perhaps there is more to it than this, but even if preclusion is not the full story, it is hard to see how – without entirely abandoning the attempt to interpret μ as some sort of generalized probability measure – one could avoid the implication that events of measure zero do not happen. If this is correct, then acceptance of the preclusion principle is a minimal requirement for reconceiving quantum mechanics along lines suggested by the path-integral formalism.

The problem then is that, thanks to interference, there are far too many sets of measure zero, so many in fact that events which are in reality able to occur seem to be ruled out as a logical consequence of the preclusion of other events that overlap them. (Remember that event = subset of Ω.) Here I'm referring to the numerous "logical paradoxes" of quantum theory, including the Kochen–Specker paradox,[2] the GHZ paradox, the Hardy paradox, and the "three-slit paradox" that I'll focus on in a moment. Each of these can be realized in terms of sets of particle trajectories together with appropriate combinations of slits or Stern–Gerlach-like devices (as in [13] or [14], for example), so that the relevant quantal measure can be discerned. What then makes all these paradoxes paradoxical is that all or part of the history space Ω is covered by precluded events. In the Kochen–Specker setup, these overlapping preclusions cover the whole of Ω, implying, according to our customary way of reasoning, that nothing at all can happen (cf. [15]). The other examples are similar, but not quite as dramatic.

The contradictions in question can be illustrated with a diffraction experiment involving not two, but three "slits." Consider, then, an idealized arrangement as shown in Figure 15.1, with source S, apertures a, b, c, and a designated location d to which the particle in question might or might not travel. (The letter d is meant to suggest "detector," but modeling one explicitly would complicate our setup

[2] This comes in two versions, the original version referring to a single spin-1 system, and the version of Allen Stairs [12] referring to an entangled pair of such systems. The latter seems to have been the first example of an obstruction to locally causal theories based purely on logic (as opposed to probability-based obstructions like the Bell inequalities).

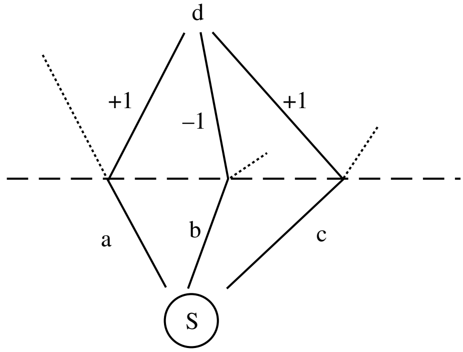


Figure 15.1 The 3-slit paradox.

unnecessarily, without changing anything essential, as long as we can assume that the detector would function properly.)

To this setup belongs a history space Ω consisting of the various possible particle trajectories, and a quantal measure μ assigning a non-negative real number to each set of trajectories. Let a be the event that the trajectory passes through slit a and similarly for b and c, and let d be the event that it arrives at d. Consider further the event A that the particle arrives at d after traversing a. (Notice here that a, b, and c are all intrinsic events, not measurement events. We are not placing detectors at any of the slits, either explicitly or implicitly.) Writing the intersection $X \cap Y$ of two arbitrary events X and Y simply as their product XY, we then have that

$$A = ad, \;\; B = bd, \;\; C = cd, \;\; d = A + B + C,$$

where in the last equation a plus sign has been used to denote the union of disjoint subsets.

Now imagine the region d to be small enough that we can represent the path integrals for A, B, and C by single amplitudes whose squares yield the (un-normalized) measures of the corresponding events, and suppose further that these amplitudes are $+1$ for events A and C and -1 for B. Then $\mu(d) = \mu(A+B+C) = |1-1+1|^2 = 1$, whereas

$$\mu(A+B) = \mu(B+C) = |1-1|^2 = 0 \; .$$

Therefore, the events $A+B = d(a+b)$ and $B+C$ are precluded even though $A+B+C$ is not and can sometimes happen.

If we think classically, this is an outright contradiction. Suppose we look for the particle at d and find it there. We can then infer that since it didn't pass through a or b ($A+B$ being precluded), it must have arrived via c. But reasoning symmetrically, we can infer by the same token that it must have arrived via a. Obviously, the two conclusions contradict each other.

In the language of coevents, we can express the situation in formulas as

$$\phi(A+B)=0,\ \phi(C+B)=0,\ \phi(d)=\phi(A+B+C)=1,$$

where the first two equations follow from the preclusion rule and the third expresses that the particle did arrive at d. Formally a contradiction can be derived from these three equations by Boolean manipulations following the classical rules of inference. If one asks which rules were used (see the Appendix), one comes up with the following list, where A and B represent arbitrary events and $\neg A = \Omega \backslash A$ is the complementary event to A.

From $\phi(A)=\phi(B)=1$ conclude $\phi(AB)=1$.
From $\phi(A)\neq 1$ conclude $\phi(A)=0$.
From $\phi(A)=0$ conclude $\phi(\neg A)=1$.
From $A\subseteq B$ and $\phi(A)=1$ conclude $\phi(B)=1$.

These formal relationships are instructive, but one can also see the root of the inconsistency informally in a way that indicates how one might think to escape it. What the formal rules really express is the ingrained belief that reality is described by a *single* trajectory γ such that an event A happens (the corresponding predicate is true) if and only if $\gamma \in A$. We might therefore be able to extricate ourselves from the contradiction if reality were given not by a single trajectory, but by some more subtle combination of trajectories, for which – in some sense – both A and C could happen simultaneously, or alternatively for which $A+C$ could happen without either A or C happening separately. We will see that the so-called "multiplicative scheme" realizes the latter alternative.

15.5 Freeing the coevent

Which came first, the history or the event? To the extent that an event is nothing but a set of histories, it might seem that the history came first, and this would be in accord with the classical worldview, where only a single history is in some sense actual.[3] On the other hand, when we consult our experience, what we meet with are events. Individual histories we experience – if at all – only as idealized limits of events. One might argue on this basis that it is the event that should come first, and this would be consistent with a more "holistic" or "dialectical" attitude toward the history space Ω (cf. category theory, toposes, etc.).

Be that as it may, the practical needs of probabilistic theories, I think, force us to accord events an independent status, for it is only they which have nontrivial

[3] I'm leaving aside here questions of "temporality": is the past "actual," or the past and the future, or only the present, or . . . ? I hope that the neglect of such questions will not unduly prejudice the rest of this discussion.

measures in general. For quantal measures this argument becomes more convincing, because the measure of an event no longer reduces, even formally, to the measures of its constituent histories. (At best, it reduces to the measures of pairs of histories.) It seems clear in particular that the concept of preclusion makes no sense at all except in relation to events. Once events are dignified in this manner, the rules governing coevents also acquire a certain freedom, and I am proposing to use this freedom in order to overcome the logical conflict between preclusion and the doctrine that reality can be described fully by a single history. More concretely, I am proposing to describe reality, not by an individual history but by an individual coevent, which mathematically is a kind of "polynomial in histories." The rules governing which coevents are dynamically possible can then change in such a way as to accommodate the preclusion principle without engendering an inconsistency. In the simplest case the polynomial is just a monomial, meaning in effect just a *subset S* of the history space. This simplest case, that of the *multiplicative scheme*, is the only one I will discuss herein. Other schemes are described in [16–18].

I am tempted at this point just to present the multiplicative scheme and discuss some of its applications, but I'm afraid that without further background, it would appear far less natural than it will if its intimate connection with deductive logic is brought out. On the other hand, I know that for some people, any hint of tampering with classical logic raises a barricade between them and the slightest sympathy with whatever comes next. For them I should emphasize that the type of scheme I am proposing can stand as a self-contained framework, whether or not one accepts a logical way of looking at it. With this caveat, let me embark on some remarks relating logic to physics that will lead in a natural way to the multiplicative scheme in the next section.

For a scientist, logical inference is – or I believe should be – a special case of dynamics. Think, for example, of forecasting the motion of Mars using Kepler's laws. Here we begin with certain events, the locations of Mars at certain earlier times, and from them we infer certain other events, namely its locations at certain later times. In other cases, we may draw conclusions from the non-occurrence of an event. Thus, from the fact that the event "sighting of the New Moon" did not happen last night, we might conclude that it will happen tomorrow night (and in consequence a new month of the Islamic calendar will begin). This second example illustrates, I hope, how inferences from dynamical laws can shade gradually into inferences from logic alone, for example the inference that if my keys are not in my pocket they must be in my jacket (which I left locked in my house). In the extreme case of simple abstract deductions like "if 'A' is true and 'B' is true then 'A and B' is also true," the inference feels so obvious that we hardly recognize it as an inference at all, but this feeling goes away for more complicated rules like

"Peirce's law":[4] Or think of the logical puzzles like, "the green house is to the right of the white house, coffee is drunk in the green house, …"

Notice here that the logic I'm speaking of concerns physical events, not strings of words and not propositions in a formal language. It is a "logic of nature," not a logic of language or thought or mathematical truth. This logic, I contend, is not prior to experience. Rather it codifies certain relations among events that, until recently, have been so ubiquitous in human experience that they have been ossified and condensed into a scheme that seems as if it were unchangeable. What we do when we predict where Mars will appear next week is exactly what the rules of logical inference do in a limited way, but also in an absolute and universal manner reminiscent of how geometry was once treated as prior to physics. But just as a better understanding of gravity forced geometry back into touch with physics, so also a better understanding of the microworld can do the same for logic. The resulting inferential scheme will not be universal but will depend (at least in part) on the particular physical system to which it is adapted. This, at any rate, will be true for the type of logic exemplified by the multiplicative scheme.

For present purposes, it helps to view a logic as a threefold structure or "triad," whose outlines tend to be obscured when formal logicians write about their subject. The first component of the triad is the event algebra $\mathfrak{A}$, a Boolean algebra whose members can be thought of as questions about the world, as described above. If we adopt this imagery then the second component of the triad is the space $\mathbb{Z}_2$ of possible answers (or "truth values"), and the third (and most neglected) is the "answering map" or coevent $\phi : \mathfrak{A} \to \mathbb{Z}_2$. In any given physical theory, $\mathfrak{A}$ and $\mathbb{Z}_2$ will be fixed but ϕ will vary in the same way that the solutions to Maxwell's equations vary. Each dynamically allowed ϕ describes then, a possible reality (or "quantal history" as one might term it). In this context, dynamical "laws of motion" in the traditional sense and rules of logical inference both take the form of conditions on ϕ. The classical rules of inference can be stated very simply if we view them in this manner. In fact, we can give them in three equivalent forms, one "deductive," one "algebraic," and one with a topological or order-theoretic flavor.[5]

Before stating these rules, however, I ought to highlight another aspect of anhomomorphic logic that removes it from the more traditional milieu. Namely, it pays equal attention to both "poles" of $\mathbb{Z}_2$, both 1 (happening, truth) and 0 (not-happening, falsehood). Whether you know that an event has happened or that it has not (as with the Moon sighting), you have learned something from which consequences can be drawn. Perhaps the reasons why falsehood has nevertheless tended

[4] $((A \to B) \to A) \to A$

[5] I have not brought the quantifiers $\forall$ and $\exists$ into the discussion because they seem to be irrelevant. One will implicitly use them in formulating the questions in $\mathfrak{A}$, but not in inferring relations among the possible answers.

to be ignored in favor of truth are first, that most logicians are mathematicians interested in deducing theorems from other propositions taken to be true; and second, that they implicitly or explicitly adopt the rule that A is true if and only if $\neg A$ is false. In the context of physics and physical events, this rule is not as self-evident as it might appear to be in a mathematical context, because answering "no" to the question "Is the particle here?" need not commit you to answering "yes" to the question "Is the particle over there?" Precisely this distinction will feature prominently in the multiplicative scheme. There is also evidence that logicians in early Buddhist times took note of it [19].

What then are the classical rules of inference expressed as conditions on ϕ? (I will assume in advance that ϕ is preclusive.) In deductive form they can, as shown by Dowker [20], be condensed into three conditions (where the symbol $\Rightarrow$ indicates deducibility):

(1a) $\phi(A) = \phi(A \to B) = 1 \;\Rightarrow\; \phi(B) = 1$ (*modus ponens*).
(1b) $\phi(A) = 0 \;\Rightarrow\; \phi(\neg A) = 1$.
(1c) $\phi(0) = 0$.

In algebraic form they are the condition:

(2) ϕ is a homomorphism of unital Boolean algebras,

which says equivalently that ϕ preserves & (*and*) and $\neg$ (*not*). And expressed as a condition on the events $\phi^{-1}(1)$ *affirmed* by ϕ, they say simply that

(3) $\phi^{-1}(1)$ is a maximal preclusive filter in Ω.

Here, following the usual definitions, a *filter* is a non-empty family Φ of events (elements of $\mathfrak{A}$, hence subsets of Ω) closed under intersection and passing to supersets. It is *preclusive* if it contains no precluded events (whence it cannot contain the empty subset 0), and it is *maximal* if it cannot be enlarged without ceasing to be preclusive. Poetically expressed, such a ϕ "maximizes being": it affirms as many events as it can, subject to fulfilling the other conditions.

In view of formulation (2), a classical coevent may be called "homomorphic," and a coevent that breaks any of the classical rules may be called *anhomomorphic*.

15.6 The multiplicative scheme: an example of anhomomorphic coevents

We have granted ourselves the freedom to change the rules ("of inference") governing coevents, but how to do so? Numerous avenues open up, but of the large number that have been explored by those of us working on the question (for some of them see [16]), only a handful have seemed promising, in the sense of permitting enough events to happen on the one hand, but restricting the coevents sufficiently to reproduce the predictive apparatus of standard quantum theory on the other hand. The current favorite seems to be the *multiplicative scheme*, which not only is among the simplest to apply, but also represents perhaps the mildest change to the classical rules. The change is so mild, in fact, that it is non-existent if we express the classical rules in the form (3) of the previous section! The difference then springs solely from the different meaning of "preclusive," or rather (because its meaning as such has not changed) from the new patterns of preclusion (patterns of precluded events) that become possible under the influence of quantal interference.

More formally, let us make the following definitions. Recall that a coevent ϕ is *preclusive* when it honors the preclusion principle, that is when $\phi(A) = 0$ for every precluded event $A \in \mathfrak{A}$. Call such a ϕ *primitive* when it follows whatever further rules of inference we have set up. The collection of all primitive coevents I will denote by $\widehat{\mathfrak{A}}$, since it is analogous in some ways to the spectrum of the event algebra $\mathfrak{A}$. The elements of $\widehat{\mathfrak{A}}$ are then (the descriptions of) the dynamically allowed "realities" or "possible worlds."

To arrive at the multiplicative scheme, we can retain condition (3) word for word as the definition of a *primitive preclusive multiplicative coevent*, or for short just "primitive coevent." Sorting out the definitions then shows that rules (1) and (2) do not survive intact. Of the first set, (1a) and (1c) survive but (1b) does not. Of condition (2), what survives is unitality and the preservation of the *and* operation (this being the origin of the name "multiplicative," since algebraically, & is multiplication). All of condition (3) survives, of course, but its meaning is probably easier to grasp when it is expressed in "dual" form.

To formulate it this way let us define first a map from sets F of histories to coevents $\phi = F^*$ by specifying that $F^*(A) = 1$ iff $A \subseteq F$. To put this in words, let's say that a coevent ϕ *affirms* an event A when $\phi(A) = 1$ and *denies* it when $\phi(A) = 0$. Our definition then says that F^* affirms precisely those events that contain F (as illustrated in Figure 15.2, where F^* affirms A but denies both B and C). When $\phi = F^*$, I will say that F is the *support* of ϕ. Now, in the case where Ω is a finite set (which we are always assuming herein), one can check that a multiplicative coevent necessarily takes the form F^* for some support $F \subseteq \Omega$. The condition (3) for primitivity then says precisely that F is as *small* as possible consistent with ϕ remaining preclusive. Given the definition of F^*, the condition for primitivity thus boils down to a rather simple criterion: the support should shrink down as much as possible without withdrawing into any precluded event.

As remarked above, "truth" or "happening" is in this context a "collective property," since it pertains to events rather than to individual histories. A multiplicative

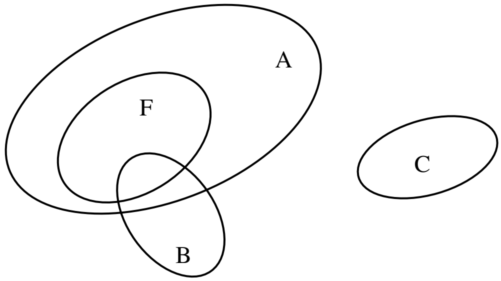


Figure 15.2 Three events and the coevent $\phi = F^*$.

coevent is also collective in nature, since it corresponds to a subset of Ω rather than an individual element.

A first test of any scheme of the present sort is that it should reproduce the classical notion of reality (namely reality as a single history) when the pattern of preclusions is itself classical. (We can take the latter to mean that an arbitrary event is precluded if and only if it is covered by precluded events. In particular, every subevent of a precluded event must be precluded.) In particular, this should happen when the quantal measure μ reduces to an ordinary measure, and also in the case of deterministic theories like classical mechanics where the dynamics reduces simply to the preclusion of an entire class of histories – those which fail to satisfy the equations of motion. It is not too difficult to verify that the multiplicative scheme passes the test in both cases. (See the theorems in Section 15.7.)

15.6.1 Resolution of the 3-slit paradox

It is a feature of the multiplicative scheme that any event E can find a primitive coevent to affirm it, as long as it is not included in some other event of zero measure. That is, there will be at least one $\phi \in \widehat{\mathfrak{A}}$ such that $\phi(E) = 1$. In our 3-slit example, the events $A + C$ and $A + B + C = d$ are both of this type, so we can see already that the multiplicative scheme will avoid the false prediction that d can never occur.

To simplify things, let's imagine that there is nothing in existence but this particular experiment and let us further ignore all histories not in d and all fine structure of the histories that are in d. Then $\Omega = d$ consists of only three elements, identifiable with the three "atoms," A, B, and C, of the event algebra $\mathfrak{A}$. With such a small history space it is easy to enumerate all the possible (multiplicative) coevents, and one sees by inspection that only two are preclusive, namely $\phi = (A + C)^*$ and $\phi = (A + B + C)^*$. The latter, however, is not primitive, since the former has smaller support. There is thus a unique primitive coevent, $\phi = (A + C)^* = A^*C^*$. With respect to this coevent, two of the eight events in $\mathfrak{A}$ happen, namely $A + C$

and d itself, and the other six do not, namely A, B, C, 0, and of course $A+B$ and $B+C$ (0 being the empty subset of Ω). In particular $\phi(d)=1$, so the paradox is removed.

This is satisfactory as far as it goes, but in one respect this 3-slit example is misleadingly simple. As the cardinality N of the history space grows, it becomes increasingly difficult in practice to work out the primitive coevents (in the multiplicative scheme there are 2^N potential supports to consider), but when the dynamics is simple enough it is possible to count them or at least to estimate their number. Typically one finds that this number also grows rapidly with N, just as one might have expected. The fact that ϕ is unique for the 3-slit setup is thus very much of an exception.

There is also another respect in which our example has been overly idealized. We have cut the experiment off at the point where the particle reaches (or does not reach, as the case may be) the location d, thereby ignoring, not only the future, but also whatever else is going on in the world besides this experiment. Both of these omissions could have serious repercussions which I'll return to briefly in the concluding section.

15.7 Preclusive separability and the "measurement problem"

Within the framework we have arrived at, individual histories are replaced in a certain sense by sets of histories while "laws of motion" are expressed via preclusion and the requirement of primitivity. In this way dynamics merges with logic to some extent, and we are able to speak directly about microscopic processes without succumbing to paradoxes of the Kochen–Specker sort – at least in simple examples. Because of its "realistic" nature, I hope that this framework will prove useful in connection with quantum gravity, specifically in the quest for a causal set of dynamics of quantal sequential growth. But a more immediate challenge is posed by the so-called "measurement problem." If the multiplicative scheme cannot solve this problem, it will be hard to take it seriously as a potential basis for unifying quantum field theory with general relativity.

Of course there is no single, well-posed "measurement problem." Rather, this phrase refers to a complex range of issues concerning the relationship of quantal processes to the macroscopic realm of classical events and "observers." Nevertheless, I think one would not be oversimplifying unduly to pose the problem as that of accounting for measurements without resorting to the notion of external agents who are not explicable in microscopic terms. In the context of the multiplicative scheme (or any of the other schemes based on anhomomorphic coevents), this problem acquires a precise formulation. One must show that the primitive coevents become classical (i.e. homomorphic) when they are restricted to a suitable subalgebra $\mathfrak{A}^{macro}$ of "instrument events."

To appreciate that this is what is needed, recall why there is a problem in the first place. Quantum mechanics as ordinarily presented either declines to describe the measurement process or it gives a manifestly false description, depending on whether or not one assumes that the state-vector "collapses" during the measurement. In the former case one is positing a phenomenon that the theory leaves in the dark, in the latter case the theory serves up a superposition of macroscopically distinct outcomes that contradicts our most elementary experiences. Now let us return to the coevent framework, where measurements are no different in principle from other quantal processes (and like other processes are to be described in terms of histories rather than evolving state-vectors). In a measurement-like situation, the theory will yield a definite set of primitive coevents to describe the different possible outcomes. For example, let events A and B be two alternative "pointer readings" in some experiment. Each of these events will correspond to a particular collection of configurations of the "pointer molecule worldlines," and will be macroscopic in the sense that the corresponding histories will involve large numbers of particles, relatively great masses, etc. If a given coevent $\phi \in \widehat{\mathfrak{A}}$ affirms A and denies B, then A is the outcome in the world described by ϕ, in the contrary case it is B. (Both types of coevent can be viable in general, since the theory is not deterministic.)[6] However, one can also construct "Schrödinger cat"-like coevents which deny both A and B, as in the 3-slit example above. Such a coevent would not be in accord with experience, which always (or almost always?) presents us with a unique outcome that does happen. Consistency with experience thus requires that no (or almost no) coevent $\phi \in \widehat{\mathfrak{A}}$ be of this ambiguous type, and this in turn is equivalent to $\phi|\mathfrak{A}^{macro}$ (the restriction of ϕ to $\mathfrak{A}^{macro}$) being classical, since when classical logic reigns, precisely one history occurs.

Formally, we can define a subalgebra $\mathfrak{A}^{macro} \subseteq \mathfrak{A}$ of macroscopic events.[7]such that disjoint elements A and B of $\mathfrak{A}^{macro}$ correspond to macroscopically distinct events in $\mathfrak{A}$. Such a subalgebra induces a partition of Ω whose equivalence classes (sets of histories distinguished by no element of $\mathfrak{A}^{macro}$) define a quotient or "coarse-graining" Ω^{macro} of Ω into "macroscopic histories." Our condition that ϕ map $\mathfrak{A}^{macro}$ homomorphically into $\mathbb{Z}_2$ is then trying to say that ϕ is supported within a single coarse-grained history (the translation being literally correct when $|\Omega| < \infty$). Put differently, the support F of such a ϕ must not overlap macroscopically

[6] Having written this however, I should add that in simple examples, the theory turns out to be much closer to deterministic than one might have expected. (See the next section.)

[7] For the events comprising $\mathfrak{A}^{macro}$ to be well defined, we might have first to condition on the happening of certain other events that are prerequisite to the existence of macroscopic objects, i.e. to the existence of what is sometimes called a "quasiclassical realm."

distinct events. When this condition is satisfied, F^* will look to $\mathfrak{A}^{macro}$ like a single coarse-grained history and will be classical in that sense.

Given the measurement problem rendered in this manner, we can solve it if we can find a sufficient condition for ϕ to behave classically and if in addition we can give reasons why the events of our macroscopic experience (almost) always fulfill the condition. To illustrate how this can work, I will quote two theorems that furnish sufficient conditions of the type we need. It should be clear from the preceding discussion that when either theorem applies, primitive coevents will behave classically as far as instrument events are concerned. One can also see the same thing by translating the conclusion of the theorems into the statement that rule (1b) above is respected. Since the multiplicative scheme validates the rest of rule (1) by construction, (1b) suffices to return us to the classical case.

The first theorem below furnishes a sufficient condition that is easy to state but more restrictive than it needs to be. The condition in the second theorem is less transparent in statement, but arguably more likely to hold in practice. (A proof of the first theorem in the finite case can be found in [16]. Proofs of the second theorem and the infinite case of the first theorem exist as well, but remain unpublished.)

Theorem 15.1 *Let $\Omega = \Omega' + \Omega''$ be a partition of Ω such that an arbitrary event $A \subseteq \Omega$ is precluded iff its intersections with Ω' and Ω'' are both precluded, and let ϕ be any primitive preclusive coevent in the multiplicative scheme. Then ϕ is supported within either Ω' or Ω''. That is, $\phi = F^*$ with $F \subseteq \Omega'$ or $F \subseteq \Omega''$.*

Theorem 15.2 *The conclusion of Theorem 15.1 persists if both Ω' and Ω'' satisfy the following weaker condition on subsets S of Ω: If any event $A \subseteq S$ lies within any precluded event B at all then it lies within a precluded event $C \subseteq S$.*

As we have seen, the important consequence of these theorems is that, when either of the respective conditions is satisfied, and with respect to any primitive coevent ϕ, either Ω' or Ω'' happens but not both. The important question then becomes whether macroscopic events are in fact "preclusively separable" in this way. This would follow immediately from the still stronger condition that: No event in Ω' interferes with any event in Ω''. However, this condition represents a very strict type of "decoherence" closely related to the idea of a *record*.[8] To the extent that one is willing to posit the existence of sufficiently permanent records of

[8] In the context of unitary quantum mechanics, the condition is satisfied for records because different versions of a given record correspond to disjoint regions in configuration space. The weaker type of decoherence usually contemplated by "decoherent historians" requires only that Ω' and Ω'' (belong to $\mathfrak{A}$ and) decohere, not that their arbitrary subevents decohere as well.

macroscopic events, one can therefore regard the measurement problem as solved. To the extent that one finds this assumption implausibly strong one can still hope to prove that macroscopic events fulfill the conditions of one of the theorems. In this sense, the measurement problem reduces to a calculation.

15.8 Open questions and further work

Before the framework presented above can be considered complete, further work will be needed on some of the questions raised by the above discussion. Foremost among these is probably the question whether one can demonstrate by examples or general arguments that the events of our macroscopic experience really are preclusively separable in the sense of the above theorems. Assuming that they are, can one explain on this basis why the textbook paradigm involving the wave function and its "collapse" works as well as it does, and if so can one quantify the deviations that one should expect from this paradigm? Here much of the way forward is clear. There exists a sketch of a derivation of the collapse rule, but it needs to be followed out in more detail. In the same direction, we of course need to recover Born's rule for probabilities, either by appeal to approximate preclusion and the Cournot principle (cf. [11]) or in some better manner. And finally, the definition of primitive coevent needs to be extended to infinite event algebras, since the most important examples of quantal dynamics (atomic physics, quantum field theory, etc.) are of this type, at least in current idealizations. Here again, there is much that could be said about work already done.

Even in its partly finished state, the coevent framework, like the Bohmian version of quantum mechanics, lets us pose questions that we would not have been able to formulate from a "Copenhagen" standpoint. Thus, for example, one can ask for the primitive coevents that describe the ground state of a hydrogen atom, or of a particle in a harmonic oscillator potential. The Bohmian particle in these cases just sits still wherever it happens to find itself. One wouldn't know by following its motion what kind of force was binding it, nor could one even know its energy in many cases. It's therefore of particular interest to ask of the multiplicative scheme, what sets of trajectories comprise the supports of the primitive coevents in these cases. Could one deduce from these sets of trajectories what the potential was and would the energy show up clearly? Currently such questions seem nearly beyond reach, in the first place because of the mathematical difficulties in defining the continuum path integral itself on a sufficiently large domain of events (cf. [21]).

More accessible, though no less interesting, are questions about experiments of the Kochen–Specker type, or about entangled pairs of particles passing through successive Stern–Gerlach analyzers. For a few gedankenexperiments of this type, people have been able to find some or all of the primitive coevents, and in some cases to study causal relationships between the coevents at earlier and later stages of the process [22, 23]. Such examples can serve as laboratories to explore possible meanings for relativistic causality, locality, and determinism within the coevent framework. For example, in a simple extension of the Hardy experiment, one finds 286 coevents ϕ of which 280 behave deterministically in the sense that the restriction of ϕ to the subalgebra of past events uniquely determines ϕ globally (cp. a similar effect found in [24]). One can also formulate different conditions of relativistic causality ("screening off") for such systems and study to what extent, and in what circumstances, they hold. If a suitable condition could be found, it could then be carried over to the causal set situation and used as a guide to formulating a quantal analog of the classical sequential growth models (cf. [14]).

If one regards the coevent schemes in logical terms, it's natural to try to bring them into relation with other non-classical logics to which anhomomorphic inference seems to bear some resemblance, such as intuitionistic, dialectical, or paraconsistent logic. With dialectics, anhomomorphic logic shares a certain tolerance of contradiction, or of what classically would be regarded as self-contradictory. With intuitionism it shares a non-classical understanding of negation, not however at the level of the event algebra, which remains strictly Boolean so that $\neg\neg A = A$ for all $A \in \mathfrak{A}$, but at the level of inference, where $\phi(A)$ becomes independent of $\phi(\neg A)$. On the other hand, whereas intuitionistic logic simply drops certain rules of inference like proof by contradiction, anhomomorphic logic adds crucially the new requirement of *primitivity*. Thus it cannot be characterized simply as either weaker or stronger than classical logic.

Returning to the coevent framework per se, I'd like to allude briefly to some of its more radical consequences, and the risks (or possibly opportunities) they hold out for this way of conceiving quantal reality. Each of these consequences is visible in simple examples. In the Hardy example referred to just above, one potentially encounters what could be called "premonitions." To be confident of their occurrence one would have to incorporate "instrument setting events" into the model, and this has not been done. Yet it looks as if the past of the coevent can determine not only events involving the particles in question, but also to some extent the settings themselves. Such an effect could be called a premonition on the part of the particles, but it could also be called a cause of the later setting event, in which case there would be no suggestion of "retro-causality." In the simple example of a particle hopping unitarily between the nodes of a two-site lattice ("two-site hopper"), one encounters a potential danger that also shows up in much the same form in connection with composite systems. In both cases it can happen that the restriction of the coevent to the subalgebra of early-time events (respectively events in one of the two subsystems) trivializes in the sense that its support becomes the whole space of partial histories. This means that only correlations between early and late times (resp. between one subsystem and the other) happen nontrivially. None of this is a problem unless carried to extremes. If, for example, the relevant time scale for this type of trivialization in realistic systems were to be comparable with the Poincaré recurrence time, then there would be little to worry about.

Finally, let me conclude with a possibility that for now is merely a dream, but which, if it came to pass, would bring with it a striking historical irony. One might discover laws that governed the pattern of preclusions without referring, directly or indirectly, to the quantal measure μ. If that happened, it would provide a more radical revision of classical dynamics (stochastic or deterministic) than that represented by the path integral. Or in the process of working out the primitive coevents in various examples, one might even discover laws expressed directly for the coevents themselves, without even needing to derive them from preclusions. If that happened, the whole superstructure of amplitudes and generalized measures would fall away, and quantum theory would have led back to something resembling classical equations of motion, but at a higher "structural" level than occupied by our old theories that identified reality with a single history.

15.9 Appendix: Formal deduction of the 3-slit contradiction

We are given the preclusions $\phi((a+b)d) = \phi((b+c)d) = 0$ and wish to deduce that $\phi(d) = 0$, assuming that ϕ follows the classical rules of inference.

Step 0: Suppose that $\phi(d) = 1$.
Step 1: If also $\phi(a+b) = 1$ then $\phi((a+b)d) = 1$, contrary to what was given. Hence $\phi(a+b) = 0$.
Step 2: $\phi(b+c) = 0$ by symmetry.
Step 3: $\phi(a+b) = 0$ implies $\phi(c) = 1$ since $c = \neg(a+b)$, the complement of $a+b$.
Step 4: From $\phi(c) = 1$ follows $\phi(b+c) = 1$, contradicting step 2.
Step 5: Therefore our supposition was false and $\phi(d) = 0$.

What conditions on ϕ did we use?

In step 1: If $\phi(A) = \phi(B) = 1$ then $\phi(AB) = 1$; if $\phi(A) \neq 1$ then $\phi(A) = 0$.
In step 3: If $\phi(A) = 0$ then $\phi(\neg A) = 1$.
In step 4: If $A \subseteq B$ and $\phi(A) = 1$ then $\phi(B) = 1$.

In the multiplicative scheme, only step 3 would fail. Notice that in reasoning *about* ϕ we have also employed classical logic, in particular proof by contradiction in steps 1 and 5.

Acknowledgments

The ideas presented above have grown out of extensive joint work with Fay Dowker, Cohl Furey, Yousef Ghazi-Tabatabai, Joe Henson, David Rideout, and Petros Wallden. Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI.

References

[1] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Spacetime as a causal set, *Phys. Rev. Lett.* **59**:521–4 (1987); R. D. Sorkin, Causal sets: discrete gravity (Notes for the Valdivia Summer School), in *Lectures on Quantum Gravity* (Series of the Centro De Estudios Cient).

[2] D. P. Rideout and R. D. Sorkin, A classical sequential growth dynamics for causal sets, *Phys. Rev. D* **61**:024002 (2000), http://arXiv.org/abs/gr-qc/9904062.

[3] R. D. Sorkin, Quantum dynamics without the wave function, *J. Phys. A: Math. Theor.* **40**: 3207–21 (2007) (http://stacks.iop.org/1751-8121/40/3207), quant-ph/0610204, http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.

[4] R. D. Sorkin, Quantum mechanics as quantum measure theory, *Mod. Phys. Lett. A* **9** (No. 33):3119–27 (1994), gr-qc/9401003, http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.

[5] R. B. Salgado, Some identities for the quantum measure and its generalizations, *Mod. Phys. Lett.* **A17**:711–28 (2002), gr-qc/9903015.

[6] R. D. Sorkin, Quantum measure theory and its interpretation, in *Quantum Classical Correspondence: Proceedings of the* 4th *Drexel Symposium on Quantum Nonintegrability*, held Philadelphia, September 8–11, 1994, edited by D. H. Feng and B.-L. Hu, pp. 229–51 (International Press, Cambridge, MA, 1997), gr-qc/9507057, http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.

[7] L. Hardy, Quantum theory from five reasonable axioms, http://arXiv.org/abs/quant-ph/0101012v4.

[8] J. B. Hartle, Spacetime quantum mechanics and the quantum mechanics of spacetime, in B. Julia and J. Zinn-Justin (eds), *Gravitation et Quantifications: Les Houches Summer School, session LVII, 1992* (Elsevier Science B.V., 1995), http://arXiv.org/abs/gr-qc/9304006.

[9] U. Sinha, C. Couteau, Z. Medendorp, I. Söllner, R. Laflamme, R. D. Sorkin, and G. Weihs, Testing Born's rule in quantum mechanics with a triple slit experiment, in *Foundations of Probability and Physics-5*, edited by L. Accardi, G. Adenier, C. Fuchs, G. Jaeger, A. Yu. Khrennikov, J. A. Larsson, S. Stenholm, *American Institute of Physics Conference Proceedings*, Vol. 1101, pp. 200–207 (New York, 2009) (e-print: arXiv: 0811.2068 [quant-ph]).

[10] Y. Ghazi-Tabatabai and P. Wallden, Dynamics & predictions in the co-event interpretation, *J. Phys. A: Math. Theor.* **42:** 235303 (2009), http://arXiv.org/abs/0901.3675.

[11] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung, von A. Kolmogoroff* (Berlin, Springer, 1933). Translated as: A. N. Kolmogorov, *Foundations of the Theory of Probability* (New York, Chelsea Publ. Co., 1956).

[12] A. Stairs, *Phil. Sci.* **50**:578 (1983); and private communication.

[13] S. Sinha and R. D. Sorkin, A sum-over-histories account of an EPR(B) experiment, *Found. Phys. Lett.* **4**:303–35 (1991).

[14] D. Craig, F. Dowker, J. Henson, S. Major, D. Rideout, and R. D. Sorkin, A Bell inequality analog in quantum measure theory, *J. Phys. A: Math. Theor.* **40:** 501–23 (2007), quant-ph/0605008, http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.
[15] S. Surya and P. Wallden, Quantum covers in quantum measure theory, http://arXiv.org/abs/0809.1951.
[16] Y. Ghazi-Tabatabai, *Quantum Measure Theory: A New Interpretation*, http://arXiv.org/abs/0906.0294 (quant-ph).
[17] R. D. Sorkin, An exercise in 'anhomomorphic logic', *J. Phys.: Conf. Ser.* **67:** 012018 (2007), a special volume edited by L. Diosi, H.-T. Elze, and G. Vitiello, and devoted to the Proceedings of the DICE-2006 meeting, held September 2006, in Piombino, Italia, arxiv quant-ph/0703276, http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.
[18] S. Gudder, An anhomomorphic logic for quantum mechanics, http://arXiv.org/abs/0910.3253 (quant-ph).
[19] R. D. Sorkin, To what type of logic does the 'tetralemma' belong?, http://arXiv.org/abs/10035735 (math:logic), http://www.perimeterinstitute.ca/personal/rsorkin/some.papers/.
[20] F. Dowker and P. Wallden, Modus ponens and the interpretation of quantum mechanics (in preparation).
[21] S. Gudder, Quantum measure and integration theory, *J. Math. Phys.* **50:** 123509 (2009), http://arXiv.org/abs/0909.2203 (quant-ph).
[22] F. Dowker and Y. Ghazi-Tabatabai, The Kochen–Specker theorem revisited in quantum measure theory, *J. Phys. A* **41:** 105301 (2008), http://arXiv.org/abs/0711.0894 (quant-ph).
[23] C. Furey and R. D. Sorkin, Anhomomorphic co-events and the Hardy thought experiment (in preparation).
[24] F. Dowker and I. Herbauts, Simulating causal collapse models, *Class. Quant. Grav.* **21:** 2963–80 (2004), http://arXiv.org/abs/quant-ph/0401075.
[25] Y. Ghazi-Tabatabai and P. Wallden, The emergence of probabilities in anhomomorphic logic, *J. Phys.: Conf. Ser.* **174:** 012054 (2009), http://arXiv.org/abs/0907.0754 (quant-ph).

16

Causal sets: discreteness without symmetry breaking

JOE HENSON

Causal sets are a discretisation of spacetime that allow the symmetries of GR to be preserved in the continuum approximation. One proposed application of causal sets is to use them as the histories in a quantum sum-over-histories, i.e. to construct a quantum theory of spacetime. It is expected by many that quantum gravity will introduce some kind of 'fuzziness', uncertainty and perhaps discreteness into spacetime, and generic effects of this fuzziness are currently being sought. Applied as a model of discrete spacetime, causal sets can be used to construct simple phenomenological models which allow us to understand some of the consequences of this general expectation.

16.1 Introduction: seeing atoms with the naked eye

At present, one of the most important tasks in theoretical physics is to understand the nature of spacetime at the Planck scale. Various indications from our current most successful theories point to this scale: quantum effects are to be expected to invalidate the general theory of relativity here. What should replace our current best understanding of spacetime? This question remains controversial as no theory of quantum gravity can yet be claimed to be complete. For example, some researchers are convinced that the kinematical structure used to replace the continuous manifolds of GR should be discrete, but others do not adhere to this requirement. George Ellis' great contribution to our understanding of spacetime, and his interest in the issue of spacetime discreteness, make this a very appropriate topic for these proceedings.

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

This situation loosely parallels a well-known debate that took place somewhat over a century ago. In that case it was not the discreteness of spacetime that was in question, but the discreteness of matter. As late as 1905 the hypothesis was still doubted by figures such as Wilhelm Ostwald and Ernst Mach. The troublesome scale in that case was around 10^{-10} m rather than 10^{-35} m, and even in those times, relevant experimental data was considerably less sparse than it is in the present case. However, it is interesting to review some of the work that went into clarifying the existence of atoms of matter. Perhaps there are some lessons to be learned for the present debate.

One of the ways to gain confidence in the atomicity of spacetime is to compare various estimates of Avogadro's number. The convergence of disparate methods of estimation, from Loschmidt's original result in 1865 to Einstein's various determinations in 1903–1905, gave a strong reason to believe that there was something to the generic hypothesis. Here we will concentrate on a method of Rayleigh's, published in 1899 [1, 2].[1] This remarkable piece of physics enabled him to estimate Avogadro's number, based on a naked-eye observation of Mount Everest. His rough but well-motivated calculations and observations enabled him to understand a fundamental aspect of the nature of matter. This stands out as a true example of elegance in physical reasoning, and an exemplar of good physics in general.

Rayleigh used his λ^4 scattering law to derive a relation between attenuation of light travelling through air and Avogadro's number. Assuming molecules to be small, spherically symmetric spheres with negligible absorption, he derived the following:

$$n = \frac{32\pi^3(\mu - 1)^2}{3\lambda^4\beta}. \tag{16.1}$$

Here, n is the number of molecules per cubic metre in air (related to Avogadro's number by $N_A = ((2.24 \times 10^{-2})n)$, μ is the index of refraction of air, assumed to be close to 1 (so that $\mu - 1 \ll 1$), λ the wavelength of the incoming light (for which Rayleigh used 600 nm) and β the scattering coefficient: light travelling through the atmosphere is extinguished by a factor of $1/e$ after travelling a length $1/\beta$. Recalling that Everest could be seen 'fairly bright' from Darjeeling, Rayleigh estimated $\beta \approx$ 160 km. The resulting value for Avogadro's number was of the correct order of magnitude, and roughly in agreement with the best estimates at the time.

This reasoning has a number of interesting features, which are relevant for the modern case. The assumption of atomicity of matter was input to the model, based on various physical arguments and expectations, rather than simply the output of some other theory. It is true that the atomic hypothesis was fairly popular by 1899,

[1] It is not entirely clear when Rayleigh developed this idea. It was formulated in response to a letter from Maxwell sent in 1873 [2]. It is not known whether Rayleigh responded during Maxwell's lifetime.

but it was not forced on Rayleigh as a derived consequence of some well-developed theory (indeed, for earlier estimations, it cannot even be claimed that the atomic hypothesis had much support at the time). Also, the calculation took place before very much was known about the dynamics of atoms. Quantum theory was still far off. The calculation depends only on a simple, basically kinematical, hypothesis about what molecules in air are like: they are small reflecting balls. However, the hypothesis was not completely generic; in order to do something useful, it was not necessary to make only the most minimal assumption of atomicity. It was only necessary that the hypothesis was simple to implement and physically well-motivated, based on the current best knowledge.

How can we apply these lessons to the problem of spacetime discreteness? The idea of building simple models to test the hypothesis of spacetime discreteness, or some aspect of it, is intriguing. Could analogous effects to the case of matter – attenuation of light, random motion, etc. – be found in this case? In the following section, a particular model for discrete spacetime is introduced, called the causal set [3]. It is then explained why this particular kinematical structure is especially physically appealing. In Section 16.3 some ideas for using this as the basis for a quantum dynamics of spacetime are presented. Finally, in Section 16.4 some mention is made of uses it can be put to in investigating the consequences of spacetime 'fuzziness'. There are several other reviews on the subject of causal sets available covering different aspects of the program [4–11], and other reviews relevant for the motivation of causal sets [12–14].

16.2 Causal sets

Spacetime discreteness is motivated by a number of arguments in quantum gravity research, for instance, the finiteness of black hole entropy which suggests a finite number of degrees of freedom living on the surface of the black hole (see [5] for a fuller list). Several approaches to quantum gravity embrace some notion of discreteness. In causal set theory, discreteness is taken as a basic hypothesis, on the strength of the physical arguments alluded to, in contrast to loop quantum gravity where discreteness is derived (not without some controversy [15, 16]) as a consequence of other assumptions [17].

Causal sets are a discretisation of the causal structure of continuum Lorentzian manifolds (meaning a differential manifold $\mathcal{M}$ with a Lorentzian metric $g_{\mu\nu}$), i.e., information about which pairs of points are in each other's light-cones, and which are spacelike-related. The points of a weakly causal[2] Lorentzian manifold, together

[2] A weakly causal Lorentzian manifold is one that contains no closed causal curves, otherwise called 'causal loops'.

with the causal relation on them, form a partially ordered set or *poset*, meaning that the set of points C and the order $\prec$ on them obey the following axioms:

(i) Transitivity: $(\forall x, y, z \in C)(x \prec y \prec z \implies x \prec z)$.
(ii) Irreflexivity: $(\forall x \in C)(x \nprec x)$.

If $x \prec y$ then we say 'x is to the past of y', and if two points of the set C are unrelated by $\prec$ we say they are spacelike (in short, all the normal 'causal' nomenclature is used for the partial order). It is an interesting fact about Lorentzian geometry that 'almost all' of the properties of a Lorentzian manifold are determined by this causal ordering. It has been proven that, given only this order information on the points, and volume information, it is possible to find the dimension, topology, differential structure, and metric of the original manifold [18, 19].

When hypothesising spacetime discreteness, a choice must be made about which aspects of the current best description remain fundamental, and which will now only be emergent. Since the causal partial order contains so much information, it is reasonable to choose this as fundamental. To achieve discreteness, the following axiom is introduced:

(iii) Local finiteness: $(\forall x, z \in C)\,(\mathbf{card}\,\{y \in C \,|\, x \prec y \prec z\} < \infty)$,

where $\mathbf{card}\,X$ is the cardinality of the set X. In other words, we have required that there only be a finite number of elements causally between any two elements in the structure (the term 'element' replaces 'point' in the discrete case). A locally finite partial order is called a causal set or *causet*, an example of which is illustrated in Figure 16.1. The *causal set hypothesis* is that the appropriate description of spacetime at the Planck scale is a causal set. Since the original results on causal

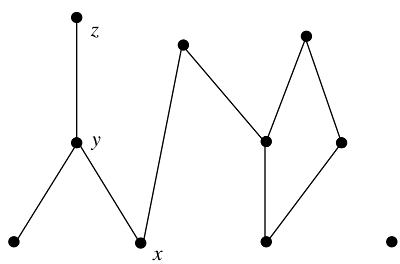


Figure 16.1 A causal set. The figure shows an example of a Hasse diagram. In such a diagram, the elements of a causal set are represented by dots, and the relations not implied by transitivity ('links') are drawn in as lines (for instance, because $x \prec y$ and $y \prec z$, there is no need to draw a line from x to z, since that relation is implied by the other two). The element at the bottom of the line is to the past of the one at the top of the line.

structure and geometry, several researchers have independently proposed this idea [3, 20, 21].

16.2.1 The continuum approximation

This gives the definition of the causal set structure itself. We now need a more precise notion of how a causal set corresponds to continuum spacetime. When can a Lorentzian manifold $(\mathcal{M}, g)$ be said to be an approximation to a causet $\mathcal{C}$? Roughly, the order corresponds to the causal order of spacetime, while the volume of a region corresponds to the number of elements representing it. But we can do a little better than this.

A causal set $\mathcal{C}$ whose elements are points in a spacetime $(\mathcal{M}, g)$, and whose order is the one induced on those points by the causal order of that spacetime, is said to be an *embedding* of $\mathcal{C}$ into $(\mathcal{M}, g)$. Not all causal sets can be embedded into all manifolds. For example, the causal set in Figure 16.2 cannot be embedded into $(1+1)$ D Minkowski space, but it *can* be embedded into $(2+1)$ D Minkowski space. Thus, given a causal set, we gain some information about the manifolds into which it could be embedded. Merely requiring that a causal set embeds into an approximating manifold is not strong enough, however. A further criterion is needed to ensure an even density of embedded elements. This relies on the concept of sprinkling.

A sprinkling is a random selection of points from a spacetime according to a Poisson process. The probability for sprinkling n elements into a region of volume V is

$$P(n) = \frac{(\rho V)^n e^{-\rho V}}{n!}. \tag{16.2}$$

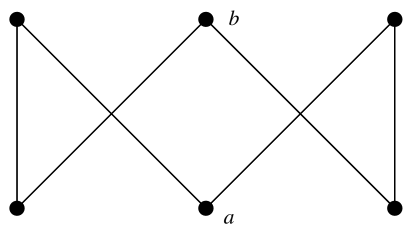


Figure 16.2 A Hasse diagram of the 'crown' causet. This causet cannot be embedded in $(1+1)$ D Minkowski space: if the above Hasse diagram is imagined as embedded into a 2D Minkowski spacetime diagram, the points at which elements a and b are embedded are not correctly related. In no such embedding can the embedded elements have the causal relations of the crown causet induced on them by the causal order of $(1+1)$ D Minkowski space. The causal set can however be embedded into $(2+1)$ D Minkowski space, where it resembles a 3-pointed crown, hence its name.

Here, ρ is a fundamental density assumed to be of Planckian order. The probability depends on nothing but the volume of the region, and so it is manifestly invariant under all volume-preserving transformations. The sprinkling also defines an embedded causal set. The Lorentzian manifold $(\mathcal{M}, g)$ is said to approximate a causet $\mathcal{C}$ *if* $\mathcal{C}$ *could have come from sprinkling* $(\mathcal{M}, g)$ *with relatively high probability*.[3] In this case $\mathcal{C}$ is said to be *faithfully embeddable* in $\mathcal{M}$. On average, ρV elements are sprinkled into a region of volume V, and fluctuations in the number are typically of order $\sqrt{\rho V}$ (a standard result from the Poisson statistics), becoming insignificant for large V. This gives the link between volume and number of elements.

It is important to note here that the way that sprinklings and embeddings are being used is to define a good discrete/continuum correspondence. That is, it is necessary to use these concepts when deciding if a Lorentzian manifold approximates to a causal set. It is not true, for example, that the fundamental structure of spacetime is supposed to be a Lorentzian manifold with a causal set embedded in it. The causal set itself is hypothesised to completely replace all continuum spacetime structures. Of course, continuum manifolds must be approximately recovered somehow, and so it is necessary to compare discrete structures to continuum manifolds in some way (this is true, explicitly or implicitly, for any discrete replacement for spacetime). This is the purpose that sprinkling serves, and it is the only way in which continua come into the story here.

Can such a simple structure really contain enough information to provide a good manifold approximation? We do not want one causal set to be well-approximated by two spacetimes that are not similar on large scales. The conjecture that this cannot happen (sometimes called the 'causal set hauptvermutung', meaning 'main conjecture') is central to the program. It is proven in the limiting case where $\rho \to \infty$ [22]. Also, all applications of the discrete/continuum correspondence have so far produced approximately unique values for important properties of continuum manifolds approximating to one casual set. This gives strong evidence for the conjecture.

As an example, consider the question of dimension. Given a fundamental causal set, how can we determine the dimension of an approximating manifold (if there is one)? It is a consequence of the hauptvermutung that all approximating manifolds have approximately the same value for this estimator. One way to answer this is to look at the proportion of points in a causal interval (otherwise known as an Alexandrov neighbourhood, the region causally between two points) that are

[3] The practical meaning of 'relatively high probability' is similar to statements about the 'typicality' of sequences of coin tosses and similar problems in probability theory. It is usually assumed that the random variable (function of the sprinkling) in question will not be wildly far from its mean in a faithfully embeddable causet. Beyond this, standard techniques involving χ^2 tests exist to test the distribution of sprinkled points for Poisson statistics.

causally related [20, 23, 24]. In the continuum, and in flat space, it is not hard to convince oneself that one half of the pairs of points in a 2D interval are related. The higher the dimension is, the smaller the proportion of related points is. Inverting this relation gives the dimension as a function of the proportion of related points:

$$D_{MM} = f^{-1}(L)\,, \tag{16.3}$$

$$f(d) = \frac{3}{2}\binom{\frac{3d}{2}}{d}^{-1}\,, \tag{16.4}$$

where D_{MM} is the *Myrheim–Meyer* dimension, and L is the proportion of points in an interval I that are related to each other. In a causal set sprinkled into this interval, the same will be approximately true. The proportion of pairs of sprinkled points that are related is $R/\binom{N}{2})$, where R is the total number of related pairs and N is the number of points sprinkled into I. Inserting this as L above gives an estimate of the dimension which is accurate when $N \gg (27/16)^{D_{MM}}$. This estimator can also be adapted to the curved-space case [24]. This gives an example of how to recover effective continuum properties from the causal set structure alone. As is clear here, the idea of sprinkling is only an intermediary used to establish the discrete/continuum correspondence, while the actual expression for the dimension only uses information intrinsic to the causal set. Similar expressions have been found to recover distances [25–27], spatial topology [28, 29], and recently the scalar curvature [30]. This gives good reason to believe that causal sets do contain enough information to pin down the continuum approximation sufficiently well.

It's important to note that some causal sets, in fact the vast majority of causal sets with a fixed large number of elements, have no continuum approximation at all [31]. In this sense the existence of the continuum is not built into causal set theory at the outset. This must be derived at a later stage (see Section 16.3 below).

16.2.2 Why this structure?

We might now choose to employ the causal set hypothesis in different ways. One use would be as a basis for a theory of quantum gravity, where the 'history space' of the quantum sum-over-histories will be made up of causal sets. A more 'Rayleighesque' use for the causal set hypothesis is to build simple but testable models which may further inform us about what is possible and what is not when it comes to spacetime discreteness. These uses are closely connected. In either case the question arises: why *this* kind of discreteness and not some other? There are many ways to discretise spacetime, after all. What singles this one out for consideration in modelling, or further, what makes it compelling as a basis for quantum gravity?

Some attractive features are already evident. Firstly there is no barrier to sprinkling into manifolds with spatial topology change, as long as it is degeneracy of the metric at a set of isolated points that enables topology change, and not the existence of closed timelike curves (one of these conditions must exist for topology change to occur, see e.g. [32] and references therein). In this discrete theory there is no problem with characterising the set of histories, as can arise in continuum path integrals. For those who believe that topology change will be necessary in quantum gravity [13, 33], this is important. Secondly, the structure can represent manifolds of any dimension – no dimension is introduced at the kinematical level, as it is in Regge-type triangulations. In fact, scale-dependent dimension and topology can be introduced with the help of course-graining, as explained in [23], giving a natural way to deal with notions of 'spacetime foam'. Also, it has been found necessary to incorporate some notion of causality at the fundamental level in other approaches to path integral quantum gravity, which hints towards using the causal set structure from the outset. But the property which truly distinguishes causal sets is local Lorentz invariance [34, 35].

It may seem contradictory to claim that a discrete structure obeys a continuum symmetry. However, this usage is carried over from condensed matter situations, and has the most physical relevance here. We might say, for instance, that a gas, liquid or glass is rotationally invariant, as opposed to, say, a crystal. This is relevant to the existence of fracture planes, and can affect propagation of sound and light in the medium. What is being referred to here is the *continuum approximation* to the discrete underlying structure rather than the structure itself. In the glass, the atomic structure does not, in and of itself, serve to pick out any preferred directions in the resulting continuum approximation.[4] Whether or not the underlying atomic configuration is rotationally invariant, or whether this question even makes sense to ask (it does not in the casual set case), is not the important point. It is also interesting to note that the large-scale symmetry is the result of randomness in the discrete/continuum correspondence in these cases.

Let us first treat the case of causal sets to which Minkowski space approximates, as in this case Lorentz symmetry is global and easier to examine. It is firmly established that causal sets are Lorentz-invariant in the above sense: whenever it makes sense to talk about global Lorentz symmetry, it is preserved. This follows from three facts [35]. Firstly, causal information is Lorentz-invariant, and so taking this from a sprinkling does not bring in any frame or direction. Secondly, the Poisson point process of sprinkling is Lorentz-invariant, as the probabilities depend only

[4] One might object that there are directions, for instance the direction from each molecule to its nearest neighbour, or some function of this for every molecule. This turns out to have no significance for any reasonable dynamics in the continuum approximation, however. And as explained below, in the case of causal sets even this weak type of direction-picking fails.

on the volume and sprinkling density. Thirdly, it has been proven that *each individual instance* of the sprinkling process is Lorentz-invariant. The proof works by showing that if each sprinkling (or each sprinkling in any measurable subset of the sprinklings) picked out a special direction, this would imply that a uniform probability measure must exist on the Lorentz group, which is impossible as the group is non-compact. No special direction can even be associated with a point in a sprinkling into Minkowski space. Therefore, since the discrete/continuum correspondence principle given by sprinkling does not allow us to pick a direction in the approximating Minkowski space, we say that causal sets are Lorentz-invariant in this sense. In other spacetimes, the existence of local Lorentz symmetry can be claimed on similar grounds.

This is arguably the intuitive outcome: random discreteness has better symmetry properties than regular discreteness, and causal information is Lorentz-invariant. A mental picture of the sprinkled points undergoing a Lorentz transformation may tempt one to believe that there is something wrong with the idea of Lorentz invariance in this context, but this is misleading. The analogy to glasses and crystals helps here. To the argument that a special direction could be picked given some small region, and so Lorentz symmetry must be broken at small scales, there is an obvious answer coming from basic special relativity: closed regions (or rulers measuring these short scales) are not Lorentz-invariant. If any physical properties can define a small region, then there is no contradiction with Lorentz symmetry if they also define a preferred direction (indeed, they must do).

Finally, phenomenological models based on causal sets manifest Lorentz invariance, helping to illustrate that the symmetry is preserved in a physically meaningful sense (see Section 16.4). This physical consideration is the root reason for calling causal sets Lorentz-invariant. Similarly, models of fields moving on lattice-like structures are well known to violate Lorentz invariance in a continuum approximation. The causal set, along with the discrete/continuum correspondence based on sprinkling, is the only known discretisation that can be shown to avoid this problem. It would be possible to take more information than the causal information from the sprinkling without violating this principle, for instance the proper lengths between sprinkled points. However, there is no reason to do this, if the causal structure and counting of points is enough to reconstruct the continuum. In this sense, the causal set seems to be a unique discretisation.

The recovery of Lorentz invariance in the continuum approximation is a remarkable property of causal sets. This is a significant advantage, whether we are interested in using the discrete structure to build simple heuristic models, or in constructing theories of quantum gravity [36]. Lattice-like structures are well known to violate Lorentz invariance in a continuum approximation, as they only allow a limited bandwidth of frequencies. Assuming Planck-scale discreteness leads to corrections to the diffusion relation, and other effects, which are currently under investigation. Bounds on these effects are now threatening to overtake the range of parameters suggested by quantum gravity-inspired Lorentz-violating scenarios [37–39]. Moreover, there is still controversy over whether introducing a Lorentz-violating cut-off in an interacting quantum theory can be made anywhere near consistent with observation, without (at best) severe fine-tuning [40–42]. If we do not find these scenarios entirely compelling, and even more so if they become ruled out, it is as well to have alternative models. Besides this, maintaining the well-established principle of local Lorentz symmetry has the advantage of severely limiting the possible types of discreteness.

16.2.3 Non-locality and causal sets

These good symmetry properties of causal sets are linked to non-locality. In lattices, each vertex is linked to a finite number of nearest neighbours, and this 'locality' makes it easy to discretise operators like the Laplacian in this context. This is not so easy for causal sets. Consider a point x in a sprinkling of Minkowski space, as illustrated in Figure 16.3. How can we define 'nearest neighbours' in a way that is intrinsic to the causal set? It could, for instance, be a sprinkled point y 'linked' to the past of x, meaning that $y \prec x$ and there are no elements z such that $y \prec z \prec x$. These linked elements must be close to x in a faithful embedding, otherwise the probability for there to be an 'in-between' element like z becomes large, as we can see from Eq. (16.2). You could imagine other definitions. Whatever they are, we have to pick

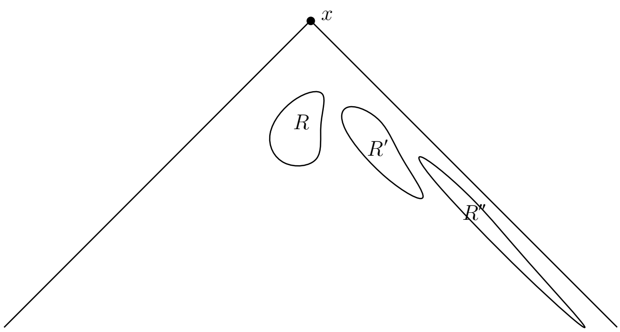


Figure 16.3 In this diagram, x is a point in a sprinkling of 2D Minkowski space, with its past light-cone. R is a region of this space, R' is a region found by boosting R, and R'' is found by boosting R' by the same factor. There are an infinite series of these disjoint regions, each of which has the same probability for containing a sprinkled element that qualifies as a nearest neighbour (by whatever Lorentz-invariant criterion). Because of this, the number of nearest neighbours must be infinite, however they are defined.

the nearest neighbours in a Lorentz-invariant way (in particular we cannot mark any other point in the spacetime). Let us try to find them to the past. Now, whatever the criterion for being a nearest neighbour, there will be some probability p for there to be one or more nearest neighbours in some finite-volume region R close to x (if the probability was 0 for any such region it wouldn't be a useful definition). But, if we boost R by a large enough factor, we can find another region R', disjoint from R. Due to the Lorentz invariance of sprinkling, the probability of finding some nearest neighbours in R' is also p. Continuing this to R'', R''', etc., we find an infinite series of regions, each with some finite probability to contain nearest neighbours. Thus, the number of nearest neighbours is infinite.

This does not come about because the information in the casual set is too sparse; it would occur even if we took more information from the sprinkling to determine our discrete structure. The argument above is quite general for any form of discreteness based on sprinklings which respects Lorentz invariance. This shows how non-locality is the flip-side of Lorentz-invariant discreteness. This isn't surprising, considering that there is an infinite volume within any finite proper time of a point in Minkowski space, due to the non-compactness of the Lorentz group. A finite number of nearest neighbours would have to lie near some particular direction away from the point x. We see in particular that the number of links from x is infinite, and linked elements occur in a narrow band that extends all the way down the light-cone, in a sprinkling of Minkowski space.

Despite this non-locality, recent work has managed to recover good approximations to local operators on causal sets by various means [30, 43], showing that this non-locality is not a terminal problem. These ideas might have interesting consequences for phenomenology, as discussed below.

16.3 Towards quantum gravity

As yet there is no quantum dynamics for causal sets that might serve as a basis for a theory of quantum gravity. However, recent progress suggests that it may soon be possible to write down such a theory and test it for consistency.

16.3.1 Growth models

There are two main approaches to causal set quantum dynamics. One, perhaps the most discussed to date, is to choose a general dynamical framework for causal set dynamics, and then constrain it by the use of some physical principles. As a stepping stone to the quantum case, a stochastic dynamics has been formulated called classical sequential growth dynamics [44, 45]. Here, the dynamical framework is one of 'growth', where causal set elements are added to the future and spacelike to existing elements, with some probability attached to each alternative way of doing so. These probabilities are then constrained to obey principles of general covariance and local causality *à la* Bell. The result is a dynamics with one free parameter per causet element, which, however, has several generic features for large classes of parameters.

This model has given rise to several lines of work. One task is to characterise the physical questions that the theory can answer, in keeping with the symmetry of general covariance, which in this context means labelling invariance. This is the causal set analog of the 'problem of time' seen in canonical theories. The problem here can be solved, in the sense that the class of meaningful questions can be found and given a simple and physically appealing interpretation [46, 47].

The model also gives some interesting results for cosmology, and the emergence of the continuum. Some generic models have a 'bouncing cosmology' with many big-bang to big-crunch cycles. The large spatial extent of the universe in these models is not a result of fine-tuning, but simply a consequence of the large age of the universe, giving a mechanism for fixing parameters that may be useful in more realistic theories [48–50]. The CSG models have also been of use in developing tests of dynamically generated causal sets to look for manifoldlike behaviour [51], and computational techniques for causal sets. It has been shown that manifoldlike causal sets are not typical outcomes for the CSG models at large scales [52]. Because of this, it seems likely that the model will remain as a useful example on the way to a quantum sequential growth model, rather than giving GR as an approximation in its own right. Recently, however, it was found that at some intermediate range of scales, some properties of CSG-generated causets were found to match the properties expected for de Sitter space [53]. The model thus continues to give some hints about what the full theory might look like.

16.3.2 Generalising the path integral

Another approach, which has until recently been less well studied, is closer to the approach taken in other quantum gravity programs such as that based on causal dynamical triangulations [54]. In this approach, a quantum dynamics is sought that is similar in form to the path integral for standard quantum theories. Roughly, one seeks to generalise the path integral for transition probabilities to the gravitational case, replacing particle trajectories or field configurations with the chosen space of discrete structures, and using an appropriate action.

Taking a more foundational view, this picture might be too limited for causal sets, or in fact any quantum gravity theory. Asking only about transition probabilities implies asking about states on spatial slices, which is not such a natural concept in a generally covariant theory; most interesting observables, for example various measures of the effective dimension for causal sets or Regge-type discretisations, are not of this type. Indeed, for a causal set very little information is given by the analog of the 'configuration on a spatial slice'. The corresponding concept for a causal set is a subset of causally unrelated elements, that is maximal, i.e. to which no further spacelike elements can be added. The only information in such a 'slice' is the number of elements it contains! This shows that the causal set is a type of discretisation in which a spacetime formulation is particularly natural, rather than a 'space and time' canonical one. To see how spacetime relates to the causal set one must look beyond a single slice; a causal set derived from sprinkling is 'non-local in time' in this sense.

A generalised framework is therefore useful. Quantum theory can of course make predictions about events that refer to properties of the history at more than one time, and all these probabilities are captured in an object called the quantum measure [55], or equivalently the decoherence functional [56] (for some recent progress on these ideas, see [57]). This object can also be given a path integral formulation for standard theories, and it is this that we seek to generalise to the gravitational case.

A toy model suggests what might be possible here. In 2D, there is an easily identifiable subclass of causal sets that contains all causal sets to which intervals of 2D Lorentzian spacetimes can approximate (and some non-manifoldlike causal sets besides) called the '2D orders'. We can make this class our history space, and supply an appropriate action which gives the continuum 2D action for causal sets that are well approximated by spacetimes (which turns out to be trivial in this case). With this done, the suprising result is that sprinklings of flat space dominate the path integral [58]. As we have seen, sprinklings were introduced to casual set theory for physical reasons, to make sure that the discrete/continuum approximation had good symmetry properties. It is remarkable and encouraging, therefore, that they also appear naturally in this setting.

However, this intriguing 2D result may not be easy to generalise. It was based on a number of theorems that are only known for the 2D orders. Also, the action is trivial in this case. The natural choice is a function of the causal set that approximates to the integral of the scalar curvature for any approximating 4D spacetime. It had until recently been difficult to construct such local quantities. Now, however, due to new work by Dionigi Benincasa and Fay Dowker, an expression for the action has been discovered [30]. The action is derived from discretisations of the d'Alembertian operator on the causal set. It uses the concept of n-element inclusive order intervals, which are subsets of the causal set comprising *all* the elements causally between some pair of elements, which have n-elements *including* that pair. A 2-element inclusive order interval is just a link. From the correspondence between volume and number, order intervals with small numbers of elements represent small regions in any corresponding spacetime, although they could be in any frame, i.e. you could find two that were highly boosted with respect to one another. The action for 4D is given by

$$\frac{1}{\hbar}S(\mathcal{C}) = N - N_1 + 9N_2 - 16N_3 + 8N_4, \tag{16.5}$$

where N is the number of elements in the causal set $\mathcal{C}$ and N_i is the number of $(i+1)$-element inclusive order intervals in $\mathcal{C}$. The lack of parameters in the action is a consequence of setting the bare Planck scale to be 1, and the bare cosmological constant Λ (which would simply add a term proportional to ΛN) to 0. The particular sequence of coefficients follows from more detailed considerations; in this respect the situation is similar to the discretisation of the second derivative for a function f on a lattice with separation Δx. A standard expression in this case is $f(x-\Delta x) - 2f(x) + f(x+\Delta x)$, where the sequence of coefficients $\{1, -2, 1\}$ is necessary to find the right continuum limit.

A remaining problem involves actually calculating Lorentzian path integrals. Writing down such an expression may now be possible, but this is not much use if no interesting expectation results can be calculated. The great strength of the CDT approach is that observables such as the scaling effective dimension, and the spatial extent of the universe as a function of time, can be calculated using Monte Carlo simulations. This is achieved by analytically continuing a parameter that multiplies the time. The path integral can then be Wick-rotated: the Lorentzian geometries are transformed to Euclidean ones, and in the process Feynman's $e^{iS/\hbar}$ turns into a more Boltzmannian $e^{-S/\hbar}$, which can then be dealt with by the well-developed methods of statistical physics.

In the case of casual sets, the idea of Wick rotation has no obvious analogue. However, analytical continuation may still be employed to calculate path integrals. In the case of a non-relativistic particle, instead of Wick rotating, the mass parameter can be continued analytically to solve the path integral, and similar methods are relevant for quantum gravity [59]. The same kinds of ideas allow a statistical path integral for causal sets to be derived from the Feynman path integral, by analytically continuing an appropriate parameter which can be included in the action. Finding some Euclideanised analogue of a causal set is therefore unnecessary. This, along with the new expression for the action, opens the door for Monte Carlo simulations in causal set theory analogous to those that have been performed in CDT theory. It would be surprising if summing over all causal sets with the action above produced a semi-classical state dominated by well-behaved 4D spacetimes; after all, this sum includes approximate spacetimes with all dimensions and topologies, and a large number of other causal sets besides these. However, this kind of work could be the beginning of a process of refinement much like that which led to CDT theory from previous models, progressing by limiting the history space or altering the action. And, as explained below, we can see that at least a 'worst case scenario' can be avoided in this dynamics, giving some hope that a well-behaved continuum approximation may be possible.

16.3.3 Problems with entropy, solutions from non-locality

As alluded to above, most causal sets do not resemble manifolds. In fact, for large number of elements N, the vast majority of causal sets have a particular and very non-manifoldlike structure [31]. A 'Kleitman–Rothschild' causal set (otherwise called a KR order in the mathematical literature) is composed of three layers, with $N/4$, $N/2$ and $N/4$ elements respectively, as shown in Figure 16.4. The elements of the top layer are to the future of some of those in the middle layer, the elements in the bottom layer are to the past of some of the elements in the middle layer, and all of those in the bottom layer are to the past of all of those in the top layer. Considering such a structure as a discrete spacetime, we have a universe with a very short duration! The number of KR orders grows like $2^{N^2/4}$ and quickly comes to vastly outnumber all others. The proportion of causal sets that have a height of greater than 3 scales like C^{-N} for some constant C [60]. Although it is hard to put a figure on the total number of manifoldlike causal sets for each value of N, it is expected that the scaling of the proportion of 'bad' configurations to 'good' ones would be even worse.

But we don't want these 'bad' histories to be the ones that dominate any path integral or 'sum over histories'. The configurations that dominate a statistical sum-over-histories arise from an interplay between the entropy of different types of history, and their actions. When the proportion of 'bad' configurations grows faster than exponentially with the number of elements or vertices in a discrete structure, this may set alarm bells ringing, for the following reason. It's tempting to think that, if the action is a sum of local terms, it should scale with the total size of the system N for a typical history, and thus the weight e^{-S} would scale exponentially with N. If that is true for bad configurations and good ones alike, we can see that

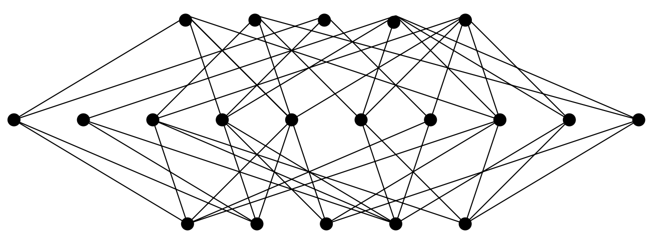


Figure 16.4 A typical Kleitman–Rothschild causal set with 20 elements. As the number of elements becomes large, this type of causal set numerically dominates all others.

the ratio between a bad and good weight can at best scale exponentially, while the proportion of bad configurations grows super-exponentially. Such an action cannot therefore prevent the bad configurations from dominating. This problem occurs, for example, if one allows all spacetime topologies in a Euclidean sum over 2D triangulations.

However, in the case of causal sets, we can see that this worry is unfounded. The action will not scale in this way for manifoldlike causal sets or for Kleitman–Rothschild causal sets.[5] This is because the number of links, and other such structures, in these kinds of causal sets do not scale like N.

Let us consider the KR orders. The first term in the action is simply N, but the second term in the action is the number of links. For any pair made up of one top-layer element and one middle-layer element, there may or may not be a link. The number of possible links between top and middle is therefore $(N/4)(N/2)$. The same is true between the bottom and middle layers, giving a total of $N^2/4$ possible links. For a typical KR order about half of these will be present, and so the second term in the action scales like N^2, not N as in the argument above.[6] It can also be seen that the other terms in Eq. (16.5) are small for a typical KR order. Taking a bottom-layer element and a top-layer element, the number of elements causally between them is the number of middle-layer elements that are linked to both. This will be around $N/8$ for a typical KR order, and so the number of 3, 4 and 5-element inclusive order intervals will scale much more slowly than N. Thus the second term dominates, and (because the sign of this term in the action will remain positive after our analytic continuation procedure, to make sure any integrals are bounded[7]) will suppress typical KR orders by a term of the form $e^{-N^2/4}$. Thus, the non-locality of causal sets is a positive feature in this case, as the action can do more to suppress bad configurations than a strictly local action could in standard cases.

This does not by itself imply that manifoldlike causal sets will dominate, but it does dispatch the argument that the faster-than-exponential growth in the non-manifoldlike configurations is an insurmountable problem. It is not yet known exactly how the action scales for a typical sprinkling into flat space; the average over sprinklings is known to be 0 (neglecting boundary terms), which is encouraging, but the size of the fluctuations also has to be taken into account. This could be settled by simulations of sprinklings. At least, we do have the 2D result to show

[5] It is interesting to note that for any graph-based discreteness, in which manifoldlike graphs have a low average valency that does not scale with N, the former is not true for some action based on graph-local quantities. It is however still true that the number of edges in a typical random graph, which would be non-manifoldlike, scales faster than N.

[6] This argument treats the elements as if they are distinguishable, or labelled, which we do not really want if we are respecting the discrete analogue of general covariance. However, it can be shown that in almost all KR orders each element is indeed distinguished from all others by the pattern of relations, i.e. they have no non-trivial automorphisms.

[7] More details of this process will be given in future publications.

that, in some statistical sums, sprinklings can arise naturally. A resolution to the question of whether they can be selected by a full sum over causal sets awaits the implementation of Monte Carlo simulations.

16.4 Consequences of spacetime discreteness

Having discussed some more ambitious uses for causal sets, we can now return to our original theme: developing simple models based on physically appealing hypotheses, and then testing them against observation and experiment. We might first consider generalising standard models, simply by replacing Minkowski space with a corresponding causal set. This helps to test if causal set discreteness was consistent with present observations, and also to look for possible new effects. Using the causal set hypothesis in this way fulfils some goals brought up in the introductory discussion. As in Rayleigh's case, this approach does not depend on the form of the dynamics for causal sets; it is essentially a non-quantum modification, which, like the atomic hypothesis for matter, might reveal some interesting consequences for discreteness before the quantum dynamics is known. Also, like the 100-year-old case, these kinds of models posit discreteness at the outset, as a hypothesis to be tested, rather than waiting for it to be derived from some fully satisfying theory. It is an input from physical considerations, rather than an output from mathematical considerations. The physical considerations in this case are the arguments pointing to discreteness, coupled with the requirement of Lorentz symmetry. Models based on causal sets therefore contrast with most work on the phenomenology of quantum gravity, which concentrates on the possibility of Lorentz violation. This alone would give good motivation to study such models.

Finally, we assume something more than the most generic possible hypothesis, that is, that quantum gravity will lead to some type of fuzziness, uncertainty or discreteness in spacetime, which will show up at or around the Planck scale. On the one hand, the more generic the hypothesis, the more significant the result might be for various approaches to quantum gravity. On the other hand, some specificity is usually necessary to actually construct a model, as Rayleigh's case illustrates. In any case, such specific models can lead to generic conclusions in one way: by providing counter-examples to generic claims. And of course any positive observations would be just as significant as for more generic models, if not more so.

One remarkable case from causal set theory is the successful prediction – the only such prediction to date from quantum gravity research – of the order of magnitude of the cosmological constant using a heuristic argument based on the possible features of a causal set quantum dynamics [10, 61, 62]. This was a genuine prediction, in the sense that it was made at a time when most researchers took Λ to be 0. This reasoning gives rise to a scenario in which the cosmological 'constant' is subject to fluctuations, something that could give rise to further predictions.

Another idea that has been investigated is a spacetime analogue of Brownian motion. Arguments have been put forward that suggest that particles moving on a causal set background may not follow geodesics, but instead might deviate from them slightly as a result of spacetime discreteness [34]. The resulting model allows a type of Lorentz-invariant diffusion of the energy–momentum of particles.[8] The magnitude of the corresponding 'diffusion constant' has not been derived from fundamental considerations; indeed, the model is only heuristic and contains assumptions that might not be true in a more detailed dynamics. However, the model is useful in that it can be tested. Bounds can be set on the diffusion constant by various laboratory experiments and astronomical observations [34, 63]. Similar models introduce Lorentz-invariant diffusion into the motion of massless particles, and consider the case of polarisation, comparing results to observations of the CMBR [64, 65].

Another line of work concentrates on putting field theory onto a background causal set. The idea is to have a field theory, described intrinsically in terms of the causal set, that approximates to the standard case when the causal set corresponds to a continuum spacetime. As mentioned above, the discretisation of the d'Alembertian operator has been achieved, allowing a classical scalar field dynamics to be defined [30, 43]. Also, significant steps towards a quantum field theory on a causal set background have been taken. Most important amongst these are the definition of Green's functions for massive scalar fields [66], including a definition of the Feynman propagator on the causal set [67]. Work to fully define such a field theory is in progress. This would not only give an opportunity to search for new effects, but also provide a new, Lorentz-invariant way of imposing a cut-off on quantum field theory.

There is another way to treat fields moving on a causal set background, that is simple enough to describe briefly here.[9] This will shed some light on the issue of the coherence of light from distant sources.

16.4.1 A discrete model of wave propagation

It has been claimed on fairly general grounds that any fuzzing in spacetime due to quantum gravity effects will lead to a loss of coherence of light from distant sources. If we replace Minkowski space with a causal set derived from sprinkling that

[8] This point has given rise to some confusion in the literature. The fact that the model does not respect momentum or energy conservation does not imply a breaking of Poincaré symmetry. This is because the usual assumptions of the Noether theorem are not respected in this case – the underlying dynamics is not based on a local Lagrangian, and the resulting process is diffusive. Adding some kind of energy bath to the model could conceivably bring it back into the Lagrangian framework, but in this case, the lack of energy conservation in the subsystem outside the bath is no surprise either.

[9] This method seems not to give rise to momentum diffusion effects.

spacetime, we do introduce uncertainties in spacetime properties, albeit only on the kinematical level; lengths between points, for example, can only be reconstructed to some finite accuracy. Would a model of propagation of light on this spacetime therefore fall victim to this problem? Using certain assumptions, some authors have derived effects that have already been ruled out by observation [68]; others argue against these assumptions [69–72], but also find in their models that some types of spacetime uncertainty are ruled out. Looking closely at the assumptions made in either case, we can see that the uncertainty is of a Lorentz-violating form: uncertainties are considered separately for the wavelength and frequencies of light. Therefore we can already see that these arguments would not apply to a causal set model. But the question of whether causal set discreteness is consistent with the coherence of light from distant sources is not yet answered by these considerations. This can be tested in a specific model which is briefly described here (more details can be found in [73]).

For these purposes, the essential features of the situation can be modelled with a massless scalar field propagating from a source to a detector. The source represents some distant astronomical source. In this case, the relevant dynamics of the scalar field are summarised in the retarded Green's functions, which in (3+1)D flat space is a delta function on the forward light-cone:

$$G(y,x) = \begin{cases} \frac{1}{2\pi}\delta(|y-x|^2) & \text{if } y \text{ is in the causal future of } x \\ 0 & \text{otherwise} \end{cases} \tag{16.6}$$

$$= \frac{1}{4\pi r}\delta(y^0 - x^0 - r), \tag{16.7}$$

where r is the spatial distance from x to y. In terms of G, the field produced by our source is given by

$$\phi(y) = \int_P G(y, x(s))\, q\, ds, \tag{16.8}$$

where P is the worldline of the source, q is its charge, and s is proper time along P. We can model the detector signal F as the integrated field in some small and distant detector region $\mathcal{D}$, giving

$$F = q \int_P ds \int_{\mathcal{D}} d^4y\, G(y, x(s)). \tag{16.9}$$

From the form of the Green's function, we can see that the signal is basically just a measure of the pairs of null-related points, one of which is in the source, and the other of which is to the future of it in the detector (it is the volume of such pairs in the relevant space, $\mathbb{M}^4 \times \mathbb{M}^4$). With certain approximations, the result for a point charge is

$$F \approx (1+v)\,\frac{q}{4\pi R}\,V\ , \tag{16.10}$$

where V is the spacetime volume of the detector, R is the spatial separation of source and detector, and v is the relative velocity of the source to the detector. Combining a moving negative and stationary positive source, to cancel the constant term in (16.10), gives a result that has the same form as the electromagnetic case. We can make $v = a\sin(\omega t)$ for some angular frequency ω and amplitude a (for simplicity we take the duration of detection to be smaller than the period of oscillation here, making the detector more than realistically accurate for most applications). The phase of the source is reflected in the detector output, and so observations in keeping with this model demonstrate the coherence of light from distant sources.

Now we want to discretise this model. We replace the 4D Minkowski space by a causal set, to which Minkowski space is only an approximation. According to our rule introduced above, this is a causal set generated from a typical sprinkling into Minkowski space. As per usual for discretisations of fields, the field ϕ becomes a real number on each causal set element. The main difficulty is how to discretise the Green's functions. We need a function that, in the limit, gives non-zero values only on the forward light-cone. To this end, consider the causet function

$$L(x, y) = \begin{cases} \kappa & \text{whenever } x \prec y \text{ and } \{x, y\} \text{ is a link,} \\ 0 & \text{otherwise,} \end{cases} \tag{16.11}$$

where $x, y \in \mathcal{C}$ are causet elements and κ is a normalising constant of order 1, that we set to make the correspondence to the continuum model correct. We have already seen that links to any element lie very close to the light-cone and extend all the way along it. In the limit of infinitely dense sprinkling, scaling κ appropriately, this function does indeed become a δ-function on the forward light-cone of x. The simple causal character of the Green's functions has thus made it very natural and easy to discretise them on the causal set.

With these definitions in place, we can proceed to redo the continuum calculation. In that case, the signal output was proportional to a volume measure of all pairs of points, one in the source and the other in the detector, that were future-null related. In the causal set case, this simply becomes a count of future directed links between the source and the detector. With suitable definitions of the source and detector region, the calculation can be performed. For a typical sprinkling into Minkowski space, the result turns out to be the same as the continuum one, with a small discrepancy overwhelmingly due to the random fluctuations inherent to sprinkling. One coherent source, which has been mentioned in the context of loss of coherence due to quantum gravity effects, is Active Galactic Nucleus PKS 1413+135, which is at a distance of order one gigaparsec. With this distance, the discrepancy between the continuum and causal set models turns out to be about one part in 10^{12}, even for unrealistically stringent parameters, and some modelling assumptions that tend to increase fluctuations. The coherence of light from distant sources is thus preserved in this model.

The conclusion here is that causal set discreteness does not cause problems for the coherence of light from distant sources. This implies that the generic idea of introducing Planck-scale 'uncertainties' or 'fluctuations' of some otherwise unspecified type does not inevitably lead to loss of coherence. This justifies the initial expectation based on considerations of Lorentz invariance. This is good in one way, but disappointing in another; the ideal situation is to make a prediction of a small discrepancy from standard theory that can then be searched for, rather than one many orders of magnitude below what can be detected. That is another reason for pursuing the models of energy and momentum diffusion mentioned above, and for more detailed consideration of the 'fluctuating cosmological constant' scenario.

16.5 Conclusion: back to the rough ground

This chapter started with a look at how a simple physical hypothesis about discreteness can, with some thought, be developed into a model that gives considerable insight. In the 19th century, the idea of atomic matter developed from a natural speculation to a more compelling proposition, based on hints from the best theories of the day. Arguments were put forward that combined understanding from these theories (like electromagnetism) with simple hypotheses (such as that atoms were hard or perfectly reflecting spheres) to develop testable models. Early on, some researchers criticised the introduction of such simplistic hypotheses,[10] which seemed to them to be unjustified shots in the dark. Instead, they urged that the atomicity of matter should follow from some more elegant generalisation of the current theories. The most notable case is Kelvin's 'vortex atom' theory [74], in which atoms were vortices in a hypothesised perfect liquid – a subject that was well-studied at this time, and undergoing many interesting mathematical advances. However, the more basic hypotheses, leading as they did to clear physical consequences (as in Rayleigh's case), turned out to have a much greater influence on the formation of the new theory.

All this does not prove that the same will be true in the case of spacetime discreteness. But, at least, it motivates the application of the same approach in this case. Because of the difficulty of obtaining relevant data, efforts to formulate a discrete

[10] Kelvin scorns the 'assumption of infinitely strong and infinitely rigid pieces of matter' as 'monstrous' in a paper on the vortex theory of matter [74].

model with clear phenomenological relevance are likely to be more challenging than they were previously. Despite this, there are several approaches to probing quantum gravity and spacetime discreteness with the aid of heuristic arguments, and beyond this, to setting up a theory of quantum gravity based on the hypothesis of discrete spacetime.

In this spirit, spacetime discreteness has been investigated above. The fact that geometry can be almost totally described in terms of the causal relation motivates the idea that it could be the causal structure that is most fundamental, and survives discretisation (rather than, for example, distance relations). Coupling the discreteness hypothesis with another physically motivated requirement, that the symmetries of GR be preserved in the continuum approximation, gives a related way to arrive at the causal set hypothesis.

There are several open avenues of research in causal set theory. There are always more 'kinematical' questions to be answered, relating to the discrete/continuum correspondence: can we, for example, approximately deduce the metric of a spatial hypersurface in an approximating manifold, by considering only information intrinsic to the corresponding causal set? And how large are the fluctuations in the new estimator for the scalar curvature, in flat space for example? After the recent success in defining the Feynman propagator for scalar fields living on a causal set, the problem of fully defining a scalar quantum field theory is of great interest and is currently under study. Defining a dynamics for vector and spinor fields would also be useful for phenomenological models. As for the dynamics of causal sets themselves, the calculation of some 'discretised path integrals' is also now within reach, as reviewed above, prompting the development of Monte Carlo simulations. This means that, for the first time, a quantum dynamics for causal sets could be studied, and, if simulations are practical, some physically interesting results could be derived from it. Also on the question of dynamics, work on the idea of sequential growth is continuing, with some studies of a possible quantum generalisation of the CSG models.

For phenomenology, further study of the heuristic derivation of the value of the cosmological constant is also an interesting area for more research. Along with this, there is the idea of looking at the consequences of the causal set action, which, due to causal set non-locality, would lead to some small corrections to the standard expression in the continuum approximation [30]. There are still many avenues to be explored in finding the consequences of fields moving on a causal set background: for instance, to see if any new effects follow from the new discretisation of the d'Alembertian operator. Similar work could also proceed in the quantum case, once the relevant theory is established. With these techniques there is some hope of making contact with data from astronomical observations, particularly with cosmology, the area that currently seems most promising as a provider of relevant data for quantum gravity research. These lines of research hopefully bring us closer to the goal of 'seeing' the discreteness of spacetime in our best current observations, as Rayleigh managed to see the discreteness of matter with the naked eye.

Acknowledgements

The author is grateful to the organisers and participants of the Foundations of Space and Time conference, celebrating George Ellis' birthday, which was a lively and interesting meeting. Thanks are also due to participants in the Causal Sets '09 conference at DIAS, Dublin, where some recent developments in the subject were discussed, and to Peter Pesic for some useful information about Rayleigh's observation.

References

[1] Rayleigh, On the transmission of light through an atmosphere containing small particles in suspension, in *Scientific Papers by Lord Rayleigh*, Vol. 4, pp. 247–405, New York: Dover, 1899/1964.
[2] P. Pesic, *Eur. J. Phys.* **26**, 183 (2005).
[3] L. Bombelli, J.-H. Lee, D. Meyer and R. Sorkin, *Phys. Rev. Lett.* **59**, 521 (1987).
[4] P. Wallden (2010), 1001.4041.
[5] J. Henson, The causal set approach to quantum gravity, in *Approaches to Quantum Gravity: Towards a New Understanding of Space and Time*, edited by D. Oriti, Cambridge University Press, 2006, gr-qc/0601121.
[6] F. Dowker, *Contemp. Phys.* **47**, 1 (2006).
[7] F. Dowker, *AIP Conf. Proc.* **861**, 79 (2006).
[8] F. Dowker (2005), gr-qc/0508109.
[9] R. D. Sorkin, (2003), gr-qc/0309009.
[10] R. D. Sorkin, First steps with causal sets, in *Proceedings of the Ninth Italian Conference on General Relativity and Gravitational Physics, Capri, Italy, September 1990*, edited by R. Cianci *et al.*, pp. 68–90, World Scientific, Singapore, 1991.
[11] R. D. Sorkin, Space-time and causal sets, in *Relativity and Gravitation: Classical and Quantum, Proceedings of the SILARG VII Conference, Cocoyoc, Mexico, December 1990*, edited by J. C. D'Olivo *et al.*, pp. 150–73, World Scientific, Singapore, 1991.
[12] R. D. Sorkin, *Stud. Hist. Philos. Mod. Phys.* **36**, 291 (2005), hep-th/0504037.
[13] R. D. Sorkin, *Int. J. Theor. Phys.* **36**, 2759 (1997), gr-qc/9706002.
[14] R. D. Sorkin (1989), gr-qc/9511063.
[15] B. Dittrich and T. Thiemann, *J. Math. Phys.* **50**, 012503 (2009), 0708.1721.
[16] C. Rovelli (2007), 0708.2481.
[17] H. Sahlmann (2010), 1001.4188.
[18] S. W. Hawking, A. R. King and P. J. McCarthy, *J. Math. Phys.* **17**, 174 (1976).
[19] D. B. Malament, *J. Math. Phys.* **18**, 1399 (1977).
[20] J. Myrheim, CERN-TH-2538.
[21] G. 't Hooft, Talk given at 8th Conference on General Relativity and Gravitation, Waterloo, Canada, August 7–12, 1977.
[22] L. Bombelli and D. A. Meyer, *Phys. Lett.* **A141**, 226 (1989).
[23] D. Meyer, PhD thesis, MIT, 1988.
[24] D. D. Reid, *Phys. Rev.* **D67**, 024034 (2003), gr-qc/0207103.

[25] G. Brightwell and R. Gregory, *Phys. Rev. Lett.* **66**, 260 (1991).
[26] R. Ilie, G. B. Thompson and D. D. Reid (2005), gr-qc/0512073.
[27] D. Rideout and P. Wallden, *Class. Quant. Grav.* **26**, 155013 (2009), 0810.1768.
[28] S. Major, D. Rideout and S. Surya, *J. Math. Phys.* **48**, 032501 (2007), gr-qc/0604124.
[29] S. Major, D. Rideout and S. Surya, *Class. Quant. Grav.* **26**, 175008 (2009), 0902.0434.
[30] D. M. T. Benincasa and F. Dowker (2010), 1001.2725.
[31] D. Kleitman and B. Rothschild, *Trans. Amer. Math. Society* **205**, 205 (1975).
[32] A. Borde, H. F. Dowker, R. S. Garcia, R. D. Sorkin and S. Surya, *Class. Quant. Grav.* **16**, 3457 (1999), gr-qc/9901063.
[33] S. W. Hawking, *Nucl. Phys.* **B144**, 349 (1978).
[34] F. Dowker, J. Henson and R.D. Sorkin, *Mod. Phys. Lett.* **A19**, 1829 (2004), gr-qc/0311055.
[35] L. Bombelli, J. Henson and R. D. Sorkin (2006), gr-qc/0605006.
[36] J. Henson, *J. Phys. Conf. Ser.* **174**, 012020 (2009), 0901.4009.
[37] G. Amelino-Camelia (2008), 0806.0339.
[38] T. Jacobson, S. Liberati and D. Mattingly (2005), astro-ph/0505267.
[39] G. Gubitosi, G. Genovese, G. Amelino-Camelia and A. Melchiorri (2010), 1003.0878.
[40] J. Collins, A. Perez, D. Sudarsky, L. Urrutia and H. Vucetich, *Phys. Rev. Lett.* **93**, 191301 (2004), gr-qc/0403053.
[41] A. Perez and D. Sudarsky, *Phys. Rev. Lett.* **91**, 179101 (2003), gr-qc/0306113.
[42] M. Visser, *Phys. Rev.* **D80**, 025011 (2009), 0902.0590.
[43] R. D. Sorkin (2007), gr-qc/0703099.
[44] D. P. Rideout and R. D. Sorkin, *Phys. Rev.* **D61**, 024002 (2000), gr-qc/9904062.
[45] M. Varadarajan and D. Rideout (2005), gr-qc/0504066.
[46] G. Brightwell, H. F. Dowker, R. S. Garcia, J. Henson and R. D. Sorkin, *Phys. Rev.* **D67**, 084031 (2003), gr-qc/0210061.
[47] F. Dowker and S. Surya (2005), gr-qc/0504069.
[48] R. D. Sorkin, *Int. J. Theor. Phys.* **39**, 1731 (2000), gr-qc/0003043.
[49] X. Martin, D. O'Connor, D. P. Rideout and R. D. Sorkin, *Phys. Rev.* **D63**, 084026 (2001), gr-qc/0009063.
[50] A. Ash and P. McDonald, *J. Math. Phys.* **44**, 1666 (2003), gr-qc/0209020.
[51] D. P. Rideout and R. D. Sorkin, *Phys. Rev.* **D63**, 104011 (2001), gr-qc/0003117.
[52] G. Brightwell and N. Georgiou, *Random Struct. Algorithms* **36**, 218 (2010).
[53] M. Ahmed and D. Rideout (2009), 0909.4771.
[54] J. Ambjorn, J. Jurkiewicz and R. Loll (2009), 0906.3947.
[55] R. D. Sorkin, *Mod. Phys. Lett.* **A9**, 3119 (1994), gr-qc/9401003.
[56] J. B. Hartle, Space-time quantum mechanics and the quantum mechanics of space-time, in *Proceedings of the Les Houches Summer School on Gravitation and Quantizations, Les Houches, France, 6 July – 1 August 1992*, edited by J. Zinn-Justin and B. Julia, North-Holland, 1995, gr-qc/9304006.
[57] F. Dowker, S. Johnston and R. D. Sorkin (2010), 1002.0589.
[58] G. Brightwell, J. Henson and S. Surya, *Class. Quant. Grav.* **25**, 105025 (2008), 0706.0375.
[59] R. D. Sorkin (2009), 0911.1479.
[60] G. Brightwell, H. J. Prömel and A. Steger, *J. Comb. Theory, Ser. A* **73**, 193 (1996).
[61] M. Ahmed, S. Dodelson, P. B. Greene and R. Sorkin, *Phys. Rev*. **D69**, 103523 (2004), astro-ph/0209274.
[62] R. D. Sorkin, *Braz. J. Phys.* **35**, 280 (2005), gr-qc/0503057.
[63] N. Kaloper and D. Mattingly, *Phys. Rev.* **D74**, 106001 (2006), astro-ph/0607485.
[64] L. Philpott, F. Dowker and R. D. Sorkin, *Phys. Rev.* **D79**, 124047 (2009), 0810.5591.

[65] C. Contaldi, F. Dowker and L. Philpott (2010), 1001.4545.
[66] S. Johnston, *Class. Quant. Grav.* **25**, 202001 (2008), 0806.3083.
[67] S. Johnston, *Phys. Rev. Lett.* **103**, 180401 (2009), 0909.0944.
[68] R. Lieu and L. W. Hillman, *Astrophys.* J. **585**, L77 (2003), astro-ph/0301184.
[69] G. Amelino-Camelia (2004), gr-qc/0402009.
[70] D. H. Coule, *Class. Quant. Grav.* **20**, 3107 (2003), astro-ph/0302333.
[71] Y. J. Ng, H. van Dam and W. A. Christiansen, *Astrophys. J.* **591**, L87 (2003), astro-ph/0302372.
[72] W. A. Christiansen, Y. J. Ng and H. van Dam, *Phys. Rev. Lett.* **96**, 051301 (2006), gr-qc/0508121.
[73] F. Dowker, J. Henson and R. D. Sorkin, Discreteness and the transmission of light from distant sources, in preparation.
[74] Kelvin, *Proceedings of the Royal Society of Edinburgh* **VI**, 94 (1867), http://zapatopi.net/kelvin/papers/on_vortex_atoms.html.

17

The Big Bang, quantum gravity and black-hole information loss

ROGER PENROSE

I argue that the common idea of applying standard quantization procedures to the space-time geometry at the Big Bang to obtain a Planck-scale chaotic geometry is likely to be wrong, whilst such a quantum-geometric structure could indeed have relevance at black-hole singularities, these appearing to lead to a necessity of information loss. These issues are addressed by re-examining the basic rules of quantum theory in a gravitational context and by viewing things from the perspective of conformal cyclic cosmology, which is dependent upon the idea of conformal space-time geometry. This kind of geometry is also central to twistor theory, a subject in which significant advances have been made in recent years.

17.1 General remarks

What follows is essentially an extended summary of my actual talk, which I hope adequately conveys the gist of what I did report at the Stellenbosch meeting. I hope, also, that it can serve as a small token of the great respect that I have for George Ellis – in the honouring of his 70th birthday – both as a person and for the enormous contributions that he has made to science and to the cause of humanity.

I briefly discuss three different topics, all of which have relevance to the nature of quantum space-time geometry. The first has to do with the very framework of quantum theory in relation to Einstein's foundational *principle of equivalence*, and provides a reason for anticipating a change in the rules of quantum mechanics when superpositions of significant displacements of mass are involved. For the second, by referring to a cosmological scheme that I have been promoting for the last five

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

years – *conformal cyclic cosmology*, or CCC – I point out that if we consider the *conformal* space-time geometry (or light-cone geometry) at the Big Bang we find that it bears a striking similarity to the conformal geometry of the remote future. This raises important questions in relation to the second law of thermodynamics, black-hole information loss, and the issue of cosmological entropy. The third topic, also related to the key role of conformal, rather than metric geometry, has to do with a resurgence of interest in twistor theory. This primarily relates to high-energy physics, but it has also revived foundational questions with regard to twistor theory as a whole, particularly in connection with the regularization of infinities.

17.2 The principles of equivalence and quantum superposition

Einstein's *principle of equivalence* is a foundation stone of the general theory of relativity; the principle of *linear superposition of states* is, likewise, one of the foundation stones of quantum mechanics. I am maintaining that there is, however, a fundamental tension between the two which, in bringing these two great theories together, will not only involve a re-examination of the ultimate nature of space-time (as is a commonly expressed aspiration of quantum-gravity exponents) but, even more importantly, a re-examination of the fundamental structure of quantum mechanics.

Consider a laboratory quantum experiment, carried out on a table top, in which the Earth's gravitational field, taken to be a constant acceleration 3-vector **a**, is to be taken into account. We can adopt one of two different procedures: the Newtonian wavefunction Ψ_N simply treats the Earth's field in the standard quantum way by adding a term into the Hamiltonian, whereas for the Einsteinian wavefunction Ψ_E we imagine measurements being made with respect to a freely falling observer, so the Earth's field vanishes, and then we transform coordinates back to those of the stationary laboratory. We find[1] that these differ merely by a phase factor

$$\Psi_E = e^{i\left(\frac{1}{6}t^3\mathbf{a}\cdot\mathbf{a}+t\bar{\mathbf{x}}\cdot\mathbf{a}\right)M/\hbar}\Psi_N, \tag{17.1}$$

t being the time, M the total mass of the quantum system, and $\bar{x}$ the Newtonian position vector of the mass centre of that system. However, the phase factor involves the non-linear term $t^3\mathbf{a}.\mathbf{a}$ in the exponent, which tells us that the two vacua differ. This is just the $c \to \infty$ limit of the Unruh effect,[2] where we find that even though the Unruh temperature goes to zero, the two vacua remain different. This would not cause any difficulties in a single background gravitational field, and indeed, well-known experiments [1], [2] performed in 1975 and later demonstrated agreement

[1] See ex[21.6] on p. 499 in Penrose [3]; also Penrose [4].
[2] B. S. Kay, private communication.

between the two perspectives. But suppose that we consider a *superposition of two gravitational fields*. This would be the case for a desk-top experiment involving the superposition of a massive object in two different locations, where the tiny gravitational field of the object in each location would be slightly different and, in the quantum superposition of the two locations, the superposition of these two gravitational fields must also be considered. Strictly, superpositions of different vacua are illegal, leading to divergent scalar products, but I argue, by reference to Heisenberg's time–energy uncertainty principle,[3] that this illegality would not show up until around a time scale, $\hbar / E_G$ where E_G is the *gravitational self-energy* of the *difference* between the gravitational fields of the two different locations of the massive object (the Earth's own field cancelling out in this calculation).

A decay time of this nature, for the spontaneous reduction of a quantum superposition, was suggested by Diosi [5]–[7], independently by myself [8], and I subsequently motivated this from the principle of general covariance [9], [10]. I came upon the present argument more recently [4]. The idea would be that, in any *measurement process* whatever – whether it be in random motions in the environment or in the deliberate movements of material in a detector – it is this gravitationally induced spontaneous reduction of the quantum state that is involved.

17.3 Cosmology and the 2nd law

The second law of thermodynamics (henceforth, the 2nd law) tells us that the Big Bang must have been a state of extraordinarily low entropy, compared with what it might have been, this specialness apparently being fully expressed in the fact that the *gravitational* degrees of freedom – as described by *Weyl* curvature – were hugely suppressed in the initial state.[4] This contrasts strikingly with the expected generic behaviour of a black-hole singularity, with wildly diverging Weyl curvature, and it is possible to attribute the very presence of a 2nd law in our own universe to this very remarkable time-asymmetric singularity structure. The *Weyl curvature hypothesis*, or WCH, is an attempted characterization of this distinction. A geometrically clearer proposal for expressing this characterization of the Big Bang has been proposed by Tod [11], according to which the *conformal* geometry of the universe is taken to extend smoothly to a Lorentzian 4-manifold *prior* to the spacelike 3-surface that conformally represents the Big Bang.

This extension of the conformal geometry of space-time is taken to be just a mathematical trick for a neat formulation of a convenient form of WCH, but we may argue [12] that since, near the Big Bang, energies get so great that particles

[3] See p.853 and note 30.37 on p.867 in Penrose [3]; also Penrose [9].
[4] See, for example, Penrose [3], 28.8.

become effectively massless, and massless particles require only the 9 numbers per point that define the conformal (i.e. null-cone) space-time structure and not the 10th, defining the metric scale, then the actual whereabouts of the Big-Bang 3-surface is of little concern for the universe's contents. The point of view [4], [13], [14] of CCC is that we *do* attach a physical meaning to this pre-Big Bang conformal manifold, identifying it with the remote future of a universe phase prior to our Big Bang, where its conformal infinity – *spacelike* because of a positive cosmological constant[5] Λ – is identified with the conformally expanded-out Big-Bang 3-surface. According to this proposal (repeated indefinitely), I refer to what we have hitherto considered to be the entire history of our universe, from Big Bang to the exponentially expanded infinite future, as the current *aeon* in an unending succession of aeons, where the 'big bang'[6] of each is identified with the infinite future of the previous one. These joins are taken to be completely smooth, as continuations of *conformal* space-time, the conformal singularities of the entire structure occurring only at black-hole singularities.

Although CCC requires this conformal matching to be precise, there is a serious conundrum thrown up by this issue *irrespective* of whether or not CCC is to be believed, which is that although the entropy of the universe (or our present aeon, if we are adopting CCC) seems to be vastly increasing, the very early universe and the very remote future appear to be uncomfortably similar to one another, according to the usage of that word in geometry, namely that the distinction between the two seems to be essentially just an enormous scale change. Moreover, a mere change of scale is essentially *irrelevant* to measures of *entropy* – where that quantity is defined by Boltzmann's formula $S = k \log V$, the quantity V being the volume of the relevant coarse-graining region in phase space – because phase space volumes are unaltered by conformal scale changes.

Do we need to take into account cosmic inflation? In CCC, inflation does not take place, or, rather, it is displaced into the aeon prior to ours and represents the ultimate exponential expansion of that aeon. But inflation does not help us with our conundrum anyway because we can simply consider the situation just *following* the cosmic moment – at around 10^{-32} s – when inflation is considered just to have turned off. The entropy subsequently gets much larger through gravitational clumping, vast entropy increases occurring when black holes are formed. The entropy increases further when Hawking evaporation takes over, and each hole (presumably) finally disappears with a rather tame little pop. Our conundrum is to understand how these apparent facts are to square with one another.

[5] I here assume that we indeed do have a positive cosmological constant, although I do realize that alternative interpretations of the data are possible; see Ellis [15].

[6] The capitalized Big Bang is reserved for the particular event that initiated our own aeon.

The answer provided by CCC is contained in the so-called information paradox of black-hole evaporation. Many physicists have tried to argue (including Hawking, with his change of heart announced at GRG17, Dublin, in 2004) that the information is somehow retrieved in subtle correlations in the Hawking radiation. I find these arguments difficult to sustain, and I believe that there is a powerful reason to believe that vast quantities of information are actually destroyed at the singularity, compare also Braunstein and Pati [16]. The consequent loss of unitarity is completely in order, in my opinion, in view of the case being made in Section 17.2, above, that unitarity must be violated in any case, by the spontaneous reduction of the wave-function (measurement) when gravitational interactions are involved. The loss of information in black holes is really a loss of a large number of degrees of freedom, which results in an enormous thinning down of the phase space. Once a black hole has evaporated away, we find that the zero of the entropy measure must be reset, because of this loss of degrees of freedom, which means, in effect, that a large number gets *subtracted* from the entropy value – simply a *renormalization* of the entropy definition. This enables the 2nd law to be fully respected throughout the entire process. (This is perhaps a subtle point, and it can only occur if there is information loss, but it provides a completely consistent – and, I believe, necessary – picture, when information loss occurs.) This gives us a resolution of the above conundrum.

Some physicists have argued, alternatively, that the ultimate maximum entropy achieved by our universe will arise not from clumping to black holes, but from the *cosmological* event horizon, which comes about from regarding such a horizon to be analogous to a black-hole horizon. This entropy turns out to be *exactly*[7] $3\pi/\Lambda$, in Planck units, where I am assuming that $\Lambda > 0$ is a cosmological *constant*. It is just a number, having nothing to do with any dynamical processes, so I am suspicious of it, and I am even more suspicious of the cosmological temperature associated with this entropy, which from the Unruh perspective should not be felt by an unaccelerated observer [17]. Moreover, it is unclear what region of the universe this entropy refers to, as it appears to be possible to violate the 2nd law if we take the region to refer to the part of the universe within the horizon itself, if this contains very large black holes, and it can certainly be violated if it refers to the universe as a whole. In addition, it is not connected with any information loss at a singularity.

Finally – and again we need not actually *adopt* CCC, that proposal merely highlighting an issue we must confront irrespective of CCC – let us ask why the Big Bang seems to be such a smooth structure, from the conformal point of view. None of the chaotic possibilities that can arise in gravitational collapse appear to have been initially present (in time-reversed form) at our Big Bang. Such things would have been initial white-hole singularities, leading to a mess of bifurcating white

[7] This follows from arguments of conformal infinity; see Penrose and Rindler [18].

holes. One should have imagined that since the radii of space-time curvatures at the Big Bang would have been initially down at the Planck scale, they should have given rise to a mess of such things, as part of wild quantum fluctuations. But these potential possibilities were just *not present* in the initial state. According to the point of view of CCC – and, again, irrespective of CCC – simply having Planck-scale radii of curvature would not be a signal that chaotic quantum gravity holds sway. My point is that for this we would require Planck-scale radii of *Weyl* curvature, i.e. of *conformal* curvature, this being the signal that the *gravitational field* has reached the level at which quantum-gravity effects would have become dominant, which would be the case at black-hole singularities. Planck-scale radii of *Ricci* curvature would not, of themselves, count as contributing to a non-classical description of the (conformal) space-time geometry.

17.4 Twistor theory and the regularization of infinities

Twistor theory[8] presents a different perspective on the concept of space-time, regarding it as a *secondary* notion, derivable from what is taken to be something more primitive than space-time, namely *twistor space*. Accordingly, the issue of quantum gravity, from this perspective, should eventually be addressed in the context of twistor space.

As a first approximation to the notion of a twistor space, we may think of the 5-space $\mathcal{P}N$ of light rays in Minkowski space-time $\mathcal{M}$. Any point of $\mathcal{M}$ can be identified in terms of the family of light rays through it, which provides us with a *Riemann sphere* of points in $\mathcal{P}N$. This suggests that $\mathcal{P}N$ has certain qualities of a *complex manifold* and indeed this correspondence is made more complete if we extend the notion of a light ray to a massless entity with spin and a null energy–momentum vector. When a phase angle is also incorporated, we basically get a *twistor*, and the space of such entities is a 4-dimensional complex vector space $\mathcal{M}$ referred to as *twistor space*.

One of the particular advantages of twistor theory is that it provides a very immediate way, in terms of free functions, of describing wavefunctions of free massless particles of arbitrary spin. Massless particles of a given helicity s are, indeed, described in terms of *free holomorphic functions* on $\mathcal{M}$ which have the specific homogeneity degree 2s-2. Strictly, these functions are to be thought of as representatives of *cohomology*, and then all aspects of the massless field equations and the positive-frequency nature of wavefunctions are automatically incorporated. This cohomology aspect of things actually connects with quantum non-locality [19], but this is something that has not been examined a great deal.

[8] See, for example, chapter 33 of Penrose [3].

These twistor functions provide very convenient representations of high-energy particles (which can be taken as massless to a good first approximation), and there has been considerable recent interest in using twistor methods to compute strong-interaction processes. For about 30 years, now, Andrew Hodges has (almost single-handedly) been developing the theory of *twistor diagrams* ([20], [21]–[23]), which are the twistor analogue of Feynman diagrams, but despite certain inherent advantages of these methods, they were not at first very practical, and there was little interest outside the twistor community. However, when in December 2003 Edward Witten introduced his ideas of *twistor-string theory* ([24]; influenced by work of and some results of Parke and Taylor [25] and of Nair [26]), there was a resurgence of interest in this topic (e.g. Mason and Skinner [27]; Arkani-Hamed *et al.*, [28]). This has provided a new stimulus to the twistor diagram theory, and recently Hodges has introduced some important new simplifying ideas which appear to have a key role, not only in regularizing infinities (and supplanting supersymmetry in various significant calculations), but also, in my opinion, giving new hope to a long-standing programme for fully incorporating general relativity into the twistor framework [29].

An important ingredient of twistor theory is the systematic employment of higher-dimensional *contour integration*.[9] When a quantum-field-theoretic calculation diverges, this is likely to correspond to a pinching of singularities in the twistor calculation, so that the appropriate compact contour ceases to exist. In order to circumvent difficulties, Hodges [30], [31] made a key proposal which involves modifying each scalar product Z.W, between a twistor Z and a dual twistor W, so that Z.W becomes Z.Wk, where k is some positive number, this displacing the offending singularity so that the required contour can now be found, and a finite answer to the (now k-dependent) integral $I(k)$ obtained. Finally, the integral

$$\int_0^\infty e^{-k} I(k) dk \tag{17.2}$$

is performed, to eliminate k. This procedure appears to have other applications. In particular, by providing a different insight into twistor contour integrals, it offers a new ingredient to a proposal that I made in 2000 for resolving a key obstruction (the googly problem) to expressing general relativity fully within the twistor programme [29].

[9] As an alternative, Witten [24], had sought to side-step such integrals by adopting the trick of pretending that space-time has the split signature (2,2), for which the twistor space would be real so that these contour integrals become real integrals and then applying some kind of anti-Wick rotation to obtain the answers for a space-time of standard Lorentzian (1,3) signature, for which twistor space is complex. But this, in my view can be regarded as merely a provisional (albeit often useful) device.

Acknowledgments

I am grateful to numerous colleagues, particularly Paul Tod and Andrew Hodges, for helpful information, and also to NSF for support under PHY00-90091.

References

[1] Colella, R., Overhauser, A. W. and Werner, S. A. (1975) *Phys. Rev. Lett.* **34**, 1472.
[2] Bonse, U. and Wroblewski, T. (1983) *Phys. Rev. Lett.* **51**, 1401.
[3] Penrose, R. (2004) *The Road to Reality: A Complete Guide to the Laws of the Universe* (Jonathan Cape, London).
[4] Penrose, R. (2009) Black holes, quantum theory and cosmology. *J. Phys.: Conf. Ser.* **174**, 012001 (15pp) http://www.iop.org/EJ/toc/1742-6596/174/1.
[5] Disi, L. (1984) Gravitation and quantum mechanical localization of macro-objects. *Phys. Lett.* **105A**, 199–202.
[6] Disi, L. (1987) *Phys. Lett.* **120A**, 377.
[7] Disi, L. (1989) Models for universal reduction of macroscopic quantum fluctuations. *Phys. Rev.* **A40**, 1165–74.
[8] Penrose, R. (1993) Gravity and quantum mechanics. In *General Relativity and Gravitation 13*, Cordoba, Argentina. Part 1: Plenary Lectures 1992. (Eds R. J. Gleiser, C. N. Kozameh and O. M. Moreschi, Institute of Physics Publishing, Bristol & Philadelphia), pp. 179–89.
[9] Penrose, R. (1996) On gravity's role in quantum state reduction. *Gen. Rel. Grav.* **28**, 581–600.
[10] Penrose, R. (2000a) Wavefunction collapse as a real gravitational effect. In *Mathematical Physics 2000*, eds A. Fokas, T. W. B. Kibble, A. Grigouriou and B. Zegarlinski (Imperial College Press, London), pp. 266–82.
[11] Tod, K. P. (2003) Isotropic cosmological singularities: other matter models. *Class. Quant. Grav.* **20**, 521–34. doi:10.1088/0264-9381/20/3/309.
[12] Rugh, S. E. and Zinkernagel, H. (2007) Cosmology and the meaning of time (Symposium, The Socrates Spirit Section for Philosophy and the Foundations of Physics, Hellabaekgade 27, Copenhagen N, Denmark).
[13] Penrose, R. (2006) Before the big bang: an outrageous new perspective and its implications for particle physics. In EPAC 2006 Proceedings, Edinburgh, Scotland, pp. 2759–62, ed. C. R. Prior (European Physical Society Accelerator Group, EPS-AG). http://accelconf.web.cern.ch/AccelConf/e06/PAPERS/THESPA01.PDF.
[14] Penrose, R. (2008) Causality, quantum theory and cosmology. In *On Space and Time*, ed. S. Majid (Cambridge University Press, Cambridge), pp. 141–95.
[15] Ellis, G. F. R. (2009) Dark energy and inhomogeneity (paper for NEB XIII conference, Thessaloniki, 2008).
[16] Braunstein, S. L. and Pati, A. K. (2007) Quantum information cannot be completely hidden in correlations: implications for the black-hole information paradox. *Phys. Rev. Lett.* **98**, 080502.
[17] Unruh, W. G. (1976) Notes on black hole evaporation. *Phys. Rev.* **D14**, 870.
[18] Penrose, R. and Rindler, W. (1986) *Spinors and Space-Time*, Vol. 2: *Spinor and Twistor Methods in Space-Time Geometry* (Cambridge University Press, Cambridge).
[19] Penrose, R. (2005) The twistor approach to space-time structures. In *100 Years of Relativity; Space-time Structure: Einstein and Beyond*, ed. A. Ashtekar (World Scientific, Singapore).

[20] Penrose, R. and MacCallum, M. A. H. (1972) Twistor theory: an approach to the quantization of fields and space-time. *Phys. Repts.* **6C**, 241–315.

[21] Hodges, A. P. (1982) Twistor diagrams. *Physica* **114A**, 157–75.

[22] Hodges, A. P. (1985) A twistor approach to the regularization of divergences. *Proc. Roy. Soc. London* **A397**, 341–74. Mass eigenstates in twistor theory, ibid, 375–96.

[23] Hodges, A. P. (1998) The twistor diagram programme. In *The Geometric Universe; Science, Geometry, and the Work of Roger Penrose*, eds S. A. Huggett, L. J. Mason, K. P. Tod, S. T. Tsou, and N. M. J. Woodhouse (Oxford University Press, Oxford).

[24] Witten, E. (2004) Perturbative gauge theory as a string theory in twistor space. *Commun. Math. Phys.* **252**, 189–258. arXiv:hep-th/0312171v2.

[25] Parke, S. and Taylor, T. (1986) An amplitude for N gluon scatterings. *Phys. Rev. Lett.* **56**, 2459.

[26] Nair, V. (1988) A current algebra for some gauge theory amplitudes. *Phys. Lett.* **B214**, 215.

[27] Mason, L. J. and Skinner, D. (2009) Scattering amplitudes and BCFW recursion in twistor space. arXiv:0903.2083v3.

[28] Arkani-Hamed, N., Cachazo, F., Cheung, C., and Kaplan, J. (2009) The s-matrix in twistor space. arXiv:0903.2110v2 [hep-th].

[29] Penrose, R. (2000a) On extracting the googly information. *Twistor Newsletter* **45**, 1–24.

[30] Hodges, A. P. (2006a) Twistor diagrams for all tree amplitudes in gauge theory: a helicity-independent formalism. arXiv:hep-th/0512336v2.

[31] Hodges, A. P. (2006b) Scattering amplitudes for eight gauge fields. arXiv:hep-th/0603101v1.

18

Conversations in string theory

AMANDA WELTMAN, JEFF MURUGAN & GEORGE F. R. ELLIS

In this closing chapter we thought it only fitting to reintroduce Thanu Padmanabhan's Hypothetically Alert Relativist Open to Logical Discussions (Harold) who, having engaged much with the popular media and with his classical relativist background, has some probing questions about string theory. Making her debut here as his correspondent is Steph, a 'String Theorist of Endless Patience and some Humility'.

Harold: It seems to me that there has been an *enormous* amount of resources spent on the various quantum gravity programs. Is there an actual *proof* that gravity has to be quantized at all?
Steph: Well, that depends on what you think would constitute a proof. Does classical mechanics have to be quantized? Apparently. Do we have a proof to that effect? No. We have a theory, it makes predictions and seems to agree with nature so we accept it. By the nature of the whole enterprise though, if we encounter a prediction that is wrong then we have to give up the theory. So, does gravity *need* to be quantized? I don't know. What I do know is that it is a fundamental interaction. The other three fundamental interactions all seem to have consistent quantum descriptions and, personally, I find that three out of four having quantum descriptions and only one being completely classical is unappealing. But maybe this is just the way nature is.
Harold: Where does the need to quantize gravity come from except for a belief in unification which may or may not be satisfied in reality?
Steph: I would say that it comes from a belief that at sufficiently small scales all the interactions exhibit quantum behaviour. This is, of course, an extrapolation but it seems to have worked till now. Unification on the other hand comes from the

Foundations of Space and Time: Reflections on Quantum Gravity, eds Jeff Murugan, Amanda Weltman and George F. R. Ellis. Published by Cambridge University Press.

realization that at the scales at which we expect that quantum effects in gravity would manifest, all the other interactions have been fully quantized already and, more importantly, as quantum theories, there seems to be no meaningful decoupling limit at the energy scales in which the gravitational interaction can be separated off from the others.

Harold: So maybe all we actually have is a quantum theory on a classical (unquantized) spacetime. This is after all how all quantum theory calculations are done, *including* the quantum gravity calculations of string theory it seems to me.

Steph: Yes and no. Yes, nearly all quantum theory computations are done in this way. Technically, this is because the key to these computations is identifying the observables in the theory (usually correlation functions) and computing these in some S-matrix scheme (via the Feynman diagrams of the theory). In this sense, we pick a background on which to perform the computations and I can see why you would be worried about the consistency of doing this in a theory of gravity, given that gravity is all about the background. But this background dependence is a fault of our computational reliance on S-matrix techniques rather than something inherently flawed with the theory. In string theory computations, what is quantized are fluctuations of the background itself and it is important to distinguish this from the fluctuations of fields in some ambient background where the fields themselves do not, in general, backreact with the background.

Harold: I don't understand then; if gravity is quantized, how is it consistent to use a theory of quantization based on a *continuum* background spacetime?

Steph: If you're quantizing the fluctuations of the background itself, then it is certainly consistent. However, I'll be the first to acknowledge that this program has not, to date, yielded anything significant. So, yes, maybe this isn't the correct way to think about what spacetime is. Personally, I am a fan of the idea that spacetime – certainly our perception of it as a coherent sequence of events – is an emergent concept, arising out of interactions. I can't make this any more precise than that at the moment though.

Harold: But gravity is expressed in spacetime curvature and so is spacetime structure. So surely, that should be quantized if gravity is quantized. Hence a procedure of quantization based on a continuum spacetime structure will be fundamentally wrong. It will not succeed in expressing the nature of a fully quantized gravitational theory.

Steph: Perhaps ... but I'm not convinced by that argument. For one, I think you are mixing the ideas of *quantum* and *discrete*. The one need not imply the other. Newtonian physics tells us to think of gravity as a force and this picture works fantastically up to certain scales. Beyond these scales this description breaks down and has to be replaced by Einstein's description of gravity as geometry – the curvature of spacetime – as you say. But I think we can both agree that even Einstein's theory seems to not apply on all scales. Certainly, if we try to apply it to singularities, like those at the centre of a black hole or the big-bang, it breaks down spectacularly. I'm not saying that any one aspect of the theory, like say the Einstein–Hilbert action, breaks down, I'm saying that it seems that the whole description of gravity as geometry seems to be failing. So perhaps in order to quantize gravity, we have to actually give up the notion altogether.

Harold: So you're asking me to give up the idea of gravity as geometry?

Steph: Well, I'm not really asking you to give up the idea of gravity as geometry. I guess I'm asking you to take seriously the idea that the geometry we see around us might be an 'emergent' property.

Harold: But then, in what sense does string theory predict the existence of gravity?

Steph: In the sense that it predicts the existence of a spin-2 field that obeys the Einstein field equations.

Harold: For a vacuum?

Steph: In fact for a non-vanishing energy momentum tensor on the right-hand side – one that is sharply constrained by the matter content of the theory.

Harold: But does it show that this spin-2 field couples to all other matter equally with the same coupling constant independent of composition, and with only positive values of the mass?

Steph: Yes.

Harold: And if it does this, does string theory also show that no other field has these same properties of coupling to all matter, so that 'gravity' is associated only with spin-2 fields, and not for example with a scalar or vector field?

Steph: Absolutely. As you may remember a scalar gauge boson cannot be the force carrier for gravity because it does not couple to the energy content of a system so that photons, for example, would not fall in a gravitational field. Observations of gravitational lensing rules this out. Vector fields, on the other hand, result in attractive *and repulsive* forces and so do not work either. All of this is actually encoded into the theory at the level of the action.

Harold: So what I'm hearing you say is that all odd-spin particles are excluded and of the even-spin gauge bosons, spin-0 particles are excluded and it's only spin-2 particles that work? What about, say, spin-4 particles?

Steph: Good question. Can I rephrase your question as 'What about higher-spin particles?'

Harold: Certainly.

Steph: OK, then let's think about how the idea of a gauge field is usually introduced. We start off with symmetries. In Yang–Mills theory, these are generated by charges which are spacetime *scalars*, λ. If we insist that these symmetries be *local* then $\lambda \to \lambda(x)$, a local function with, in general, non-vanishing gradients. These gradients generate extra terms in the action which spoil the symmetry. To compensate for these terms we need to add a spacetime *vector*, a spin-1 particle which transforms as $A_\mu \to A_\mu + \partial_\mu \lambda$. This story isn't special to Yang–Mills theories either. Supersymmetry transformations, for example, are generated by conserved spin-$\frac{1}{2}$ supercharges. To compensate for the additional terms generated by gradients of these fields in the action, a spin-$\frac{3}{2}$ gravitino must be added. In a theory like gravity where the conserved charge is an *energy–momentum vector*, a spin-1 particle, the compensating field needed to make sure that the theory remains generally coordinate-invariant is a spin-2 field, the metric tensor. Theories whose conserved charges have spin greater than 2, are so constrained that the scattering amplitudes are essentially zero. This means that the only way to include such higher-spin gauge fields is to make the theory *non-interacting*. This argument is nicely encompassed in the Coleman–Mandula theorem (and some of its generalizations).

Harold: OK, I'll buy that, *but* a theorem is only as good as its assumptions.

Steph: Sure. And this is no exception. One way around this argument is if the higher-spin fields have non-vanishing commutators. But since these commutators tend to include generators of even higher spin, building a non-trivial higher-spin gauge theory will necessarily require fields of arbitrarily large spin. These are the so-called Vasiliev theories.

Harold: Ok, that's all well and good but, what about the principle of equivalence? Can you really get the equivalence principle out of the theory? I would consider this to be *the* central principle of general relativity.

Steph: Indeed, so would I. Actually, I've always wondered how relativists and cosmologists think about the principle of equivalence.

Harold: Well, there are, of course, two statements of the principle. The strong version is the equivalence between inertial and gravitational mass but I'd be happy if you could show me how even the weak equivalence principle, the universality of free-fall, is encoded in the theory.

Steph: But, practically speaking, if I gave you a theory of gravity defined by, say, some action, how would you tell if it encodes the equivalence principle or not?

Harold: Well, I'd compute the timelike geodesics for, say, massive test particles and see if free-fall was universal. The strong equivalence principle is a little more subtle I'll admit.

Steph: But you agree, I hope, that it will come down to an appropriate definition of masses?

Harold: Certainly.

Steph: In that case, the most elegant argument I know of comes from Weinberg and goes all the way back to 1964. Let me see if I can do it some justice. We first need to define what gravitational mass is. To this end, consider a multi-particle scattering process in which a *soft graviton*, i.e. one with 4-momentum $k \to 0$, is emitted (Figure 18.1). We can then define m_g as the strength of the interaction

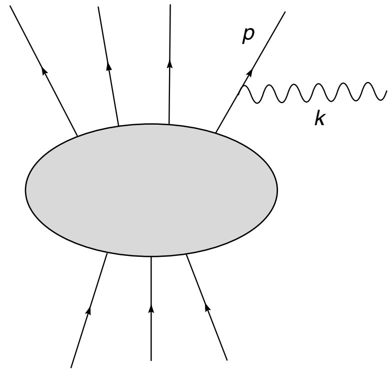


Figure 18.1 Soft graviton emission.

with the soft graviton. The amplitude for this emission $M^{\mu\nu}_{\alpha\beta}(k)$ is a product of the amplitude $M_{\alpha\beta}$ for the same process without the soft graviton emission and a kinematical factor $\sum_n \frac{\eta_n g_n p_n^\mu p_n^\nu}{p_n^\mu k_\mu - i\eta_n \varepsilon}$, that depends on the 4-momentum p_n of the nth particle and the coupling g_n between the graviton and the nth particle.

Harold: What about η_n?

Steph: It just encodes ingoing $\eta_n = -1$ or outgoing $\eta_n = +1$ particles in the scattering process. The amplitude is Lorentz-invariant only if $k_\mu M^{\mu\nu}_{\alpha\beta}(k) = 0$. This, in turn, is only true if $\sum_n g_n p_n^\mu$ is a conserved quantity. But, Weinberg reminds us, in any non-trivial process, the *only* linear combination of momenta that is conserved is the *total* momentum $\sum_n \eta_n p^\mu$. This means that all the couplings g_n must be equal and we may as well set $g_n = \sqrt{8\pi G}$.

Harold: So what you're saying is that this low-energy, massless, spin-2 particle couples to *all* forms of energy in the same way?

Steph: Precisely!

Harold: Great. I hate to be a stickler but what about the 'strong' equivalence principle?

Steph: Well, following Weinberg's argument leads to an effective gravitational mass related to the inertial mass m_i by $m_g = 2E - \frac{m_i^2}{E}$. It's easy to test this relation: for non-relativistic particles, $E \to m_i$ and $m_g = m_i$ while relativistic particles require $m_i \to 0$ and result in $m_g = 2E$. You of all people will no doubt recognize this from GR.

Harold: I certainly do. OK, I think I understand. However, while we're on the topic of 'fields', I've heard from many sources that string theory is full of them. If these fields (e.g. a scalar 'dilaton' field) are indeed there, then why do we not experience them in, for example, the solar system? Surely, the solar system tests presumably exclude them to high precision?

Steph: The dilaton in string theory has a unique interpretation as the string coupling – the strength of string–string interactions. So if you are referring to some scalar–tensor type gravitational theory, while these are not excluded from string theory I think there are quite stringent constraints put on them. The 'fields' that are usually referred to in string theory come from the geometric data of the compactified manifold and manifest in the *particle physics* of the 'large' four dimensions. This is also the reason why we don't *experience* them in everyday life.

Harold: Let me see if I understand something. String theory is a two-dimensional conformal field theory, right?

Steph: Correct, the worldsheet theory is.

Harold: But then this two-dimensional space gets embedded in a target space with ten dimensions, and six of these dimensions get compactified to give a four-dimensional 'effective' space? With the string moving in this higher-dimensional spacetime?

Steph: Yes, that is exactly what happens.

Harold: OK, so then something is bothering me. I think it goes back to the old Kaluza–Klein idea itself. Kaluza–Klein theory is effectively a fibre bundle over the four-dimensional spacetime, with a preferred projection structure that must be preserved when one makes coordinate changes; which is why general coordinate changes are not allowed (you can't mix the fifth dimension with the four). So it is not a properly covariant theory. I think this is why it was abandoned. Plus of course, the scalar degree of freedom is not observed.

Steph: I disagree. It wasn't abandoned. It morphed into the structure of modern gauge theory. Pauli's famous objection was an objection to the idea of gauge invariance as first proposed by Weyl [3]. This objection certainly stands but it is *not* how we think of gauge symmetries today (or, for that matter, anytime since Weyl's reformulation of the idea). Anyway, regarding your objection here, the point is that the five-dimensional theory *is* invariant under the five-dimensional group of diffeomorphisms. Given that this is the case, to get to the four-dimensional spacetime, we have to make some coordinate choice or in the language of gauge theory, a gauge is selected. With this choice, the five-dimensional group of diffeomorphisms splits into a four-dimensional group of diffeomorphisms and a $U(1)$ corresponding to the compatification circle. Asking for the gauged-fixed theory to be invariant under the full five-dimensional diffeomorphism group is just not meaningful: why should I expect to be able to perform a coordinate transformation and turn a photon into geometry? The four-dimensional spacetime is invariant under four-dimensional coordinate transformations as it should be, and the remaining $U(1)$ manifests in an abelian gauge theory – Maxwell eletromagnetism. I realize that an immediate objection to this is that the particular direction I chose to compactify on is arbitrary, but this is just a choice of gauge. The observables of the theory do not care whether x^1 or x^5 is the circle direction.

Harold: So what you're saying is that you just identify points in a six-dimensional submanifold of a ten-dimensional spacetime to get six small spatial dimensions. Does this not result in extra terms in the Einstein field equations?

Steph: Well, it's a little more complicated than that. I'm not really sure what you mean by 'just identify points' but if I take it literally, this would result in a 6-torus which, while a nice toy model, is not a consistent compactification. When the ten dimensions split into 6+4, the geometric and topological characteristics of the internal space manifest in the particle physics in the four dimensions. This *necessarily* means that the four dimensional space is not empty. You are forced by the consistency of the theory to have matter in the non-compact dimensions. String theory is not consistent otherwise. One *cannot* treat that problem as one does in classical general relativity by adding in matter of a specific type by hand because at the scales at which string theory is valid *everything* (matter and gravity) is quantum and coupled and there is no sensible decoupling limit in which one can consider quantum gravity alone (at least not in four dimensions).

Harold: Hmm. Actually I now think my above statement is wrong. Hence if we have a five-dimensional vacuum (or higher) and then compactify to get effective four dimensions, the four-dimensional Einstein field equations will still have extra terms arising from this process, even though one did not explicitly do an embedding. In Kaluza–Klein these extra terms are the electromagnetic and scalar field stress tensor terms in the four-dimensional Einstein field equations presumably. They will not be zero in general. Hence one will not get four-dimensional vacuum Einstein field equations in general from five-dimensional vacuum Einstein field equations.

Steph: That is correct. But I fail to see the problem: the universe we live in is certainly not a vacuum. You could argue that you should be able to satisfy solar system tests, etc. but I am no more convinced that string theory need apply to the solar system than quantum mechanics should. General relativity does and as long as I can show that general relativity is a consistent low-energy limit of string theory, I think this is OK.

Harold: Great, I am glad you mentioned this. The effective four-dimensional general relativity is supposed to be *derived* from the string theory, not just to be a consistent limit. It should be a necessary outcome.

Steph: Yes and no. Four-dimensional general relativity needs to be *the* four-dimensional, low-energy classical limit of the theory. As such, it is a necessary outcome.

Harold: This seems to be agreeing with me!! Then you can't say string theory does not apply to the solar system; the story is supposed to be that string theory is *the basis* of the theory of gravity – that is a major selling point of string theory.
Steph: In very much the same way as we can't apply quantum mechanics to the solar system even though we know that it is the underlying theory governing atomic scales, string theory as a description of nature is valid only at energy scales near the Planck scale. And just like we take a semiclassical $\hbar \to 0$ limit of quantum mechanics, to approach scales that can be probed at accessible energies, we have to average over all the quantum gravity states to obtain a low-energy effective field theory description. When we conduct solar system tests, it is this effective field theory (in this case general relativity plus matter) that we're testing.
Harold: That must mean including gravity at solar system scales. If not, where does the theory of gravity at that scale come from?
Steph: Gravity at that scale comes from the following procedure:

(1) You identify all the correct degrees of freedom in the quantum gravity theory.
(2) Integrate out all the high-energy degrees – the low-energy (infrared) physics is insensitive to these.
(3) Obtain an appropriate low-energy effective field theory – here general relativity + something close to the standard model.
(4) Make the usual approximation regarding the matter distribution in the solar system.

Harold: OK, good. So what I want to see is the averaging in step 2.
Steph: You and I both. The above procedure is an *in-principle* outline. In practice, we're more-or-less still trying to figure out step 1.
Harold: There's something else that's been bothering me. I've just read that string theory has tachyonic modes? If this is true, why is this not a problem for the theory?
Steph: Let me ask you Harold, when you say 'tachyon', what do you mean?
Harold: I mean a particle that moves faster than the speed of light, of course!
Steph: I thought so. It's what I would have expected an excellent relativist like yourself to say. Unfortunately, it's not a useful way to think of it in a quantum field theory.
Harold: So what *is* a tachyon then?
Steph: Perhaps I can illustrate with a simple example – a scalar field theory with quartic potential. The dynamics are encoded in the (symbolic) Lagrangian $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 + \mu^2\phi^2 - \lambda\phi^4$. We know already from classical mechanics that the coefficient of the quadratic term is usually interpreted as the mass of the field. The problem is that here this coefficient has the wrong sign! If we were pushed, we'd have to interpret this as a 'tachyon' since its mass-squared is negative.

Harold: That's a bit strong isn't it? The only reason you have the wrong sign is because you're writing the Lagrangian in terms of variables suited to an expansion around the origin.

Steph: Precisely! *And* this is an *unstable* point in the potential. It's a local maximum in fact. If we were to write things in terms of the real lowest energy state at $\phi = \pm\mu/\sqrt{2}\lambda$, we would find a mass term with the correct sign.

Harold: So...

Steph: So, the lesson to be learned here is that the presence of a tachyon in the spectrum of a field theory is nothing more than a signal that we're focusing on an unstable vacuum. This is exactly what happens in the 26-dimensional bosonic string and suggests that the theory is unstable. It is an instability that is *not* present in the 10-dimensional superstring theories.

Harold: Aah, wait a second! You seem to have traded the tachyonic instability for supersymmetry. Is that right?

Steph: Again ... yes, and no. Yes, in the case of the critical string theories, cancellation of the tachyonic mode in the spectrum does happen because of supersymmetry. No, because this isn't the only mechanism for achieving it. Non-critical string theories, for example, do not require supersymmetry to cancel the tachyon, at least at tree-level.

Harold: What do you mean by 'critical' vs 'non-critical' string theory?

Steph: Well, one of the main consistency conditions on string theory is that the worldsheet theory must be conformally invariant. Classically, this is not a problem but at the quantum level this symmetry can be broken, producing a *conformal anomaly*. For superstring theory, this anomaly vanishes only if the worldsheet theory consists of 10 bosons (each accompanied by a corresponding fermion). Since each boson is interpreted as a dimension of the flat target space (i.e. *spacetime*), this gives a critical dimension of 10. There is another way of getting rid of the conformal anomaly by making the target space non-trivial. The result is an anomaly-free string theory that does not have to have critical dimension. These theories are clearly of little phenomenological use but nevertheless provide excellent toy-models within which to learn about the structure of string theory.

Harold: I see. In fact this answers my next question about the number of dimensions in string theory, where they come from and how many there are in each of the string theories. However, I noticed in that argument that you say that in the critical cases, the string target space is flat. Is critical string theory *only* defined in flat spacetimes? If that's the case, it's a bit trivial isn't it?

Steph: Not at all. Flat space just happens to be one of three maximally supersymmetric 10-dimensional spacetimes on which string theory is usually studied. The other two are the 10-dimensional pp-wave and the celebrated $AdS_5 \times S^5$ spacetime. If fact, not only can we define string theory on these backgrounds; in the pp-wave we can actually quantize the string exactly while all indications are that the worldsheet theory in $AdS_5 \times S^5$ is a fully integrable theory. Both of these backgrounds also play a crucial role in the gauge theory/gravity duality.

Harold: Interesting. The dualities being discovered via string theory/M-theory sound quite amazing. The fact they exist certainly seems to imply that this theory is saying something foundational about mathematics, and is related to physics. Would it be true to say that there has been a shift in emphasis, so that string theory is now more a theory about these dualities rather than a fundamental proposal for a unified view of the foundations of physics and of spacetime?

Steph: I wouldn't necessarily say that. Certainly in the past decade, these dualities have taught us a fortune about the structure of both quantum gravity as well as gauge theories at strong coupling, things we would not have learnt easily otherwise.

Harold: Such as?

Steph: Such as minimally viscous fluids like the quark–gluon plasma, such as the idea of geometry and topology arising as emergent phenomena, such as an understanding of the microstates of black holes and even the structure of scattering amplitudes in Yang–Mills theory. In fact, I'd go so far as to say that if we do end up showing that $\mathcal{N} = 8$ supergravity is finite to all orders, it will be because of intuition gained from the gauge/gravity duality.

Harold: Come now, at least one other theory of quantum gravity purports to give a correct counting of the microstates of a black hole.

Steph: I'm not disputing that there are other ways to count microstates. However, state counting is a *kinematic* problem. In my opinion, more difficult is the issue of *dynamics*, as encoded, for example, in the greybody spectrum of the black hole. As far as I am aware, string theory is the only approach to date that allows a computation of this.

Harold: Fair enough. But you still haven't answered my question about the shift in focus of research in string theory.

Steph: I was getting there. The way I see it – and I'm quite certain that many would disagree with me on this – if anything, string theory has shifted from being one single theory to a *collection of ideas* all based on the premise that the fundamental objects in the theory are one-dimensional strings. Let me explain. Take quantum field theory. It certainly isn't one single theory; scalar field theory, quantum electrodynamics, Yang–Mills theory, the standard model and a host of other 'theories' are all specific examples of quantum field theories. They are all ideas, based on the premise that the field is the fundamental object and its excitations provide the necessary quanta. For me, string theory – the ultraviolet completion of quantum field theory – has a similar flavour. AdS/CFT, flux compactifications, string cosmology, matrix theory, D-branes and a host of other 'theories' are ideas that constitute 'string theory'. At the end of the day, each of these ideas teach us something about the nature of the whole and since the whole – string theory – is fundamentally a quantum theory of, not only gravity but all the interactions, I would say that it remains a proposal for a unified view of these interactions.

Harold: OK, then let me ask you a question that has always bothered me. What are strings made of? What is doing the vibrating? What is the substance that underlies the existence of strings, and hence of matter?

Steph: Nothing!

Harold: What do you mean?

Steph: If you ask that question in the context of string theory, then strings are *fundamental*, they have no internal structure. In this sense, strings are not at all like the classical strings you see around you. These have internal structure, they are made up of molecules and atoms. Consequently, classical strings have both transverse and longitudinal oscillation modes. The strings of string theory, however, *only* have transverse oscillations – a hallmark of the fact that they have no internal constituents. This is not unlike in particle physics, the fundamental object is a point particle. As such, we don't think of these particles as having any internal structure.

Harold: Except that we know that the correct description of particle physics is quantum field theory and these things we call particles are nothing but the quanta of the field, so in some sense it is the *field* that is fundamental.

Steph: Exactly! Something very similar exists for string theory too. It is called *string field theory* and its relation to the fundamental string is the same as that of the quantum field to the particle. But expanding on this, I suspect, will be a whole other conversation.

Harold: Does string theory really say something about cosmology? Aren't the energies involved so high that it has no real link to cosmology?

Steph: The truth is that you'll get a different answer depending on who you ask. So let me give you both my opinion and, I think, the opinion of the more conservative amongst us. String theory *may* have something to say about cosmology. And cosmology *may* be able to show us something about string theory. The fact that the energies associated with string theory are so very high tells us that the natural link to cosmology is in the very early universe. It is certainly a possibility that, with improved data resolution in the near future, we may be able to see the inner workings of a quantum gravity theory within the effects we observe. Of course whether we will be able to distinguish between the various candidate theories is another story. On the other hand, any such effects may also just be washed out by inflation [4]. The moduli of string theory also present possible candidates for the inflaton or the field responsible for dark energy. In this sense string theory may be the very fundamental theory that cosmologists have been waiting for to explain their effectively phenomenological solutions to the early and late time acceleration of the universe. Likewise cosmology may provide the ultimate testing ground for string theory and provide the opportunity to connect to observables. Some – in fact many – would say that it is too early to ask these questions of string theory. The theory is not yet completely understood so how can we expect to understand cosmology from it yet. This is a fair criticism but I don't think it will stop this very active area of research and should we hit on a real smoking gun signal of string theory in the night sky I suspect the dissenting views will change.

Harold: Let me then ask you a question about the successes and failures of the theory itself. You mention dark energy as a possible future success of the theory, but is it not a past failure? Let's talk more about dark energy or the cosmological constant or the energy of the vacuum, whatever you choose to call it. Is it not the case that string theory predicts that only non-positive values are possible? Didn't this prediction come before the surprising 1998 observation of the accelerated expansion of the universe [5] (i.e. a positive cosmological constant?) Don't the observations rule out the theory?

Steph: This is a good question because it bears on both the theory and observations. There was no prediction of a value for the cosmological constant as such before 1998. It was only during the late 1990s that string theorists were discovering the non-perturbative tools so important in understanding the full theory [6]. In particular, the discovery of dualities [7], branes [6], black-hole entropy counting [8], matrix theory [9] and the AdS/CFT duality [10] were mostly studied in the context of supersymmetry – which we very much know is not the state the universe is in today. We now do have solutions with positive cosmological constant [11, 12] but they necessarily break supersymmetry. These solutions were found within the context of solving the problem of moduli stabilization – a problem that is ripe for the ideas of cosmology. Scientifically, perhaps this misconception arose out of the prominent no-go theorems of Maldacena and Nunez [13]. As with all no-go theorems these papers make several very restrictive assumptions. Under these assumptions they state that no solutions can be found with a positive cosmological constant. However, these theorems do not result in a broad constraint on string theory in any sense. In particular, one of the constraints requires that *no localized sources* are included. Relaxing this very constraint and breaking supersymmetry allows for us to find solutions with a positive cosmological constant and these are not constraints that one would consider natural given that we know the universe is not supersymmetric at the low-energy scales (or late times) associated with dark energy and we expect a full theory to include non-perturbative ingredients. In fact there are now many proposals for the solution of the dark energy problem as well as for a fundamental theory of inflation either based directly on string theory or inspired from string theory ideas – see for example [12, 14, 15].

Harold: You mentioned string theory solutions to dark energy. Doesn't this just lead really to a multiverse and ultimately away from verifiability not towards?

Steph: It is only fair to say that string theory solutions to dark energy are still a matter of much controversy and yes, the multiverse is at the heart of this. But in fact I think it is only correct to recognize that in fact the direction that string theory may be going here is really an explanation that was proposed by Steven Weinberg in the 1980s already [16]. It is easy in retrospect to think that dark energy was always considered to be a possible outcome and that a positive cosmological constant was expected or worked on by the majority. In fact it was not. The observation came as a great surprise to the bulk of the physics and astronomy communities, though perhaps not to Steven Weinberg. In the 1980s Weinberg [16] was working on the vacuum energy problem and puzzling over why the calculated value from quantum field theory is so high when an observer would never possibly conclude that such a high value was realized as dark energy. Ultimately it turns out that the observed value is 123 orders of magnitude smaller than the calculated expected value. Weinberg had a possible solution though. He argued that if the underlying theory had multiple vacua each ultimately describing a different potential universe, then it could explain why the vacuum energy that we observe is not large but is both small and not zero. This thinking does not sound foreign to us now, but remember that this was against all conventional wisdom at the time and remarkably Weinberg's prediction that the cosmological constant would be small but positive was exactly what was observed over a decade later. Now it may be that Weinberg was at least partially correct or even completely correct, and that string theory is this underlying theory with the multiple vacua provided by flux compactifications. The multivacuum property of string theory that is so often mocked as its greatest challenge may ultimately be lauded as providing the very solution that nature requires to the cosmological constant problem – the solution glimpsed by Weinberg nearly 30 years ago.

Harold: But is the landscape an accepted part of string theory? Or is it a proposal supported by only some string theorists?

Steph: Let's start by reminding ourselves how the landscape arises in string theory. To do so we have to discuss a little about the longstanding moduli stabilization problem. Moduli are the degrees of freedom associated with a particular compactification, and in the low-energy effective action they appear as exactly massless scalar fields. You may remember that we discussed how particle physics is encoded in geometry in string theory. This is the same idea. The fact that these fields are massless reflects on the geometry side a freedom for changing the geometry of the compactification in certain ways. While there is no energy change associated with this change – on the particle physics side we are left with fields that can evolve freely in space and time. If these fields couple to all matter fields and in particular if they couple strongly or even at gravitational strength to all matter fields then equivalence principle violations will be the result. We essentially have introduced a fifth force. There are quite a few creative attempts at solutions to this problem

[11, 14, 17, 18] but it is only fair to say that this is still an area of active research. This question though does lead to the possibility of connecting to cosmology – because it is in cosmology that scalar fields are most used. The landscape arises in the context of trying to solve the moduli stabilization problem. Naively these moduli are exactly massless – i.e. they have $V = 0$ – so there is no cost to moving around in this moduli space – i.e. a flat landscape so to speak. If we now find some mechanism (for example flux compactifications [11, 19]) to give these moduli non-trivial potentials then the result is a possibly complicated landscape with each point representing a different set of values for the moduli. So first we must recognize that the landscape arises through this solution to a specific problem. It may be that we are entirely on the wrong track and that the moduli problem is solved in an entirely different way or even that the compactifications could be non-geometric and have an entirely different interpretation. I think it is only fair to say it is too early to decide if the landscape is an inevitable outcome in string theory. I think it is also worth pointing out that this landscape exists in quantum field theory as well. It is not unique to string theory. However somehow in string theory the interpretation has led to ideas of multiverses and it has opened up a Pandora's box of philosophical questions that I do not think are necessarily implied.

Harold: Is there some place in the landscape that genuinely gives the standard model of particle physics? Or is it able to produce only something standard model-like?

Steph: We don't know yet. One problem with the landscape is that it is vast. Different solutions involve turning on fluxes of various types. These fluxes are quantized and so values of the fields in the landscape are not continuously tunable – there are only discrete values allowed. Whether there is any point in the landscape that gives us the standard model is not yet known. This is very much work in progress though – see [20, 21] for details.

Harold: If string theory can explain the cosmological constant problem, you can hardly call this a prediction. Any more than finding a spin-2 particle in its spectrum means string theory predicts gravity! In fact even finding the standard model in the landscape would be impressive but not a prediction. Are there areas where string theory can still make predictions and we can still say we have experimental verifiability?

Steph: So what you are really asking is whether string theory is physics. Because ultimately the physics process is about theory and experiment each walking together, sometimes one is ahead but unless they are connected we are either mathematicians and philosophers or engineers. And the answer is a resounding yes. String theory is physics. It is rich in mathematics but all of this structure and rigour is aimed at connecting to the observable universe and understanding the puzzles left by other fields. Other than the possible connections with cosmology, there are also connections with phenomonelogy and the possibility of seeing signatures at the LHC. Perhaps even more compelling are the unexpected connections to the studies of the quark gluon plasma produced at the Brookhaven National Laboratory [22]. Physicists in both areas have found that some of the properties of this plasma can be understood using the powerful tool of string theory – duality. In fact there are some properties of the plasma, like its viscosity-to-entropy ratio, that are better modelled as a black hole in a five-dimensional space than as a clump of nucleons in the usual four dimensions. The string theory model works far better than could have been expected and, in some ways, does not yet capture some of the higher-order or imperfect effects in the laboratory. So the predictions are not as sharp as they could be but they are certainly there.

Harold: Surely this is more of a mathematical spinoff rather than an actual prediction?

Steph: Ever the sceptic! I will quote Joseph Polchinksi [23]: 'One of the repeated lessons of physics is unity. Nature uses a small number of principles in diverse ways. And so the quantum gravity that is manifesting itself in dual form at Brookhaven is likely to be the same one that operates everywhere else in the universe.' So yes Harold – while we may not yet have a conclusive smoking gun or bullet-in-the-head of string theory test, what we do have is a broad range of tantalizing hints and the hope that through the sometimes small steps and occasional giant leaps of science we will get there.

Acknowledgments

The questions of Harold mostly represent issues raised by several colleagues – in particular G. F. R. E. – in a series of email discussions (beginning in 2006) and daily coffee debates with the string theory group at UCT, mostly J. M. We are grateful to a number of people who have helped shape the discussion carried out in this chapter. In particular, we would like to thank Thanu Padmanabhan for the idea of Harold and permission to reuse him as our inquisitor. Further, we are indebited to Alex Hamilton, Julien Larena and Andrea Prinsloo for stimulating much of the coffee discussions. J. M. owes much of his understanding of string theory to conversations with Antal Jevicki, Horatiu Nastase and especially Robert de Mello Koch. Finally, we are immensely grateful to our sponsors: the NRF of South Africa and the National Institute for Theoretical Physics, the Foundational Questions Institute and Cambridge University Press.

References

[1] J. Schwinger, *Particles, Sources and Fields*, Volume I (Perseus Books, USA, 1998).
[2] C. P. Burgess, *Living Rev. Rel*. **7**:5 (2004), 2003 e-Print: gr-qc/0311082

[3] L. O'Raifeartaigh, *The Dawning of Gauge Theory* (Princeton University Press, USA, 1997).
[4] R. Easther, W. H. Kinney and H. Peiris, *JCAP* **0505**:009 (2005), 2004 e-Print: astro-ph/0412613; R. Easther, W. H. Kinney and H. Peiris, *JCAP* **0508**:001 (2005), e-Print: astro-ph/0505426.
[5] Supernova Search Team (Adam G. Riess *et al.*), *Astron. J.* **116**:1009–38 (1998), e-Print: astro-ph/9805201.
[6] J. Polchinski, *Phys. Rev. Lett*. **75**:4724–7 (1995), e-Print: hep-th/9510017.
[7] For a review, see C. Vafa, e-Print: hep-th/9702201.
[8] A. Strominger and C. Vafa, *Phys. Lett*. **B379**:99–104 (1996), e-Print: hep-th/9601029.
[9] T. Banks, W. Fischler (Texas U.), S. H. Shenker (Rutgers U., Piscataway) and L. Susskind, *Phys. Rev*. **D55**:5112–28 (1997), e-Print: hep-th/9610043.
[10] J. Maldacena, *Adv. Theor. Math. Phys*. **2**:231 (1998), e-Print: hep-th/9711200.
[11] S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, *Phys. Rev*. **D68**:046005 (2003), e-Print: hep-th/0301240.
[12] S. Kachru, R. Kallosh, A. D. Linde, J. M. Maldacena, L. P. McAllister and S. P. Trivedi, *JCAP* **0310**:013 (2003), e-Print: hep-th/0308055.
[13] J. M. Maldacena and C. Nunez, *Int. J. Mod. Phys*. **A16**:822–55 (2001), e-Print: hep-th/0007018.
[14] J. Khoury and A. Weltman, *Phys. Rev*. **D69**:044026 (2004), e-Print: astro-ph/0309411; J. Khoury and A. Weltman, *Phys. Rev. Lett*. **93**:171104 (2004), e-Print: astro-ph/0309300.
[15] P. Brax, C. van de Bruck, A. C. Davis, J. Khoury and A. Weltman, *Phys. Rev*. **D70**:123518 (2004), e-Print: astro-ph/0408415; P. Brax, C. van de Bruck, A. C. Davis, J. Khoury and A. Weltman, *AIP Conf. Proc*. **736**:105–10 (2005), e-Print: astro-ph/0410103.
[16] S. Weinberg, *Rev. Mod. Phys*. **61**:1 (1989).
[17] T. Damour and A. M. Polyakov, *Nucl. Phys*. **B423**:532–58 (1994).
[18] S. Carroll, *Phys. Rev. Lett*. **81**:3067–70 (1998), e-Print: astro-ph/9806099.
[19] R. Bousso and J. Polchinski, *JHEP* **0006**:006 (2000), e-Print: hep-th/0004134.
[20] Standard model bundles. See, for example, R. Donagi, B. A. Ovrut, T. Pantev and D. Waldram, *Adv. Theor. Math. Phys*. **5**:563–615 (2002), e-Print: math/0008010.
[21] R. Blumenhagen, M. Cvetic, P. Langacker and G. Shiu, *Ann. Rev. Nucl. Part. Sci*. **55**:71–139 (2005).
[22] P. Candelas *et al.*, *Adv. Theor. Math. Phys*. **12**:2 (2008), e-Print: hep-th/0502005.
[23] See 'On some criticisms of string theory', at http://www.itp.ucsb.edu/ joep/.

Index