Toward a theory of everything

by

Ding Jia

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Robert Oeckl Faculty Member, Centro de Ciencias Matemáticas, UNAM
Supervisor(s):	Lucien Hardy Faculty Member, Perimeter Institute for Theoretical Physics
	Achim Kempf Professor, Dept. of Applied Mathematics, University of Waterloo
Internal Member:	Robert Mann Professor, Dept. of Physics and Astronomy, University of Waterloo
Internal-External Member:	Doreen Fraser Associate Professor, Dept. of Philosophy, University of Waterloo
Other Member(s):	Lee Smolin Faculty Member, Perimeter Institute for Theoretical Physics

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A physical theory of everything is supposed to tell us: (1) The dynamical laws for matter and gravity; (2) The boundary condition of the universe; (3) The relation between the theory and experience. I present a personal assessment for different possibilities in addressing these tasks, which leads to the following picture for the most likely case.

The dynamical laws for matter and gravity are captured by a joint matter-gravity path integral. The boundary condition is such that all configurations are summed over indifferently. The interior condition selects for individual experiences and nothing else.

A human life is characterized by a sequence of experiences. Probabilistic predictions for each individual experience can be found in path integrals of the above form, but no collective account for joint experiences can be given. Individual experiences are related by their enabling conditions, which encode memories. Comprehension of objects are constructions, derived from regularities and repetitive patterns among the experiences.

Acknowledgements

This thesis is driven by the wish to know the truth of Nature.

In human societies, the pursuit of truth in itself is a luxury. I am deeply grateful to my supervisors Lucien Hardy and Achim Kempf, to colleagues, staffs, funders and donors of University of Waterloo and Perimeter Institute for making this pursuit possible for me.

Yet for humans, the pursuit of truth in itself is also a necessity. Some sacrificed their lives for it, and people hardly sacrifice lives for what is unnecessary. This thesis is impossible without those pursuers of truth, who sacrificed in one way or another. You inspired me to join the effort. Thank you.

Dedication

To Hui Jia, Li Jin, and Yicen Lu.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Objectives

This thesis is an attempt toward a physical theory of everything. In this thesis, a physical theory of everything is taken to tell us:

1. The dynamical laws for matter and gravity.
2. The boundary condition of the universe.
3. The relation between the theory and experience.

None of these is a solved problem, and the objective of the thesis is to work toward the true answers.

1.2 Strategic discussion

For each topic listed above, there are different possibilities for the true answer. It is stupid to pick an answer and hope it is true. It is wiser to look for the true answer.

One way to find out the true answer is to collect a list of possibilities, estimate their chances, and constantly update the list and credence upon new inputs. In this section I offer my list, and explain the reasons behind the credence assignments.

1.2.1 Matter

For the dynamical law of matter, the Standard Model of particle physics provides the current best candidate theory. Due to issues such as dark matter, there remains the question what theory captures the full truth of Nature. My guess for the possibilities are:

1. Standard Model as it stands, or a straightforward extension

   Credence: $> 70\%$.

A straightforward extension means an alternative theory with a modified set of matter species or/and a modified Lagrangian (e.g., adding dark matter to the Standard Model), but still within the functional integral framework.

My confidence in this possibility is high, because of the empirical success of the Standard Model and the functional integral framework.

2. Something else

   Credence: $< 30\%$.

1.2.2 Gravity

The dynamical law for gravity may be captured by a theory of quantum gravity, or something else. My guess for the possibilities are:

1. Geometric-variable path integrals

   Credence: $\sim 50\%$.

   In General Relativity, gravity is captured by spacetime geometry. Straightforwardly, path integrating over spacetime geometries yield quantum theories of gravity. Examples include versions of simplicial quantum gravity (quantum Regge calculus) and dynamical triangulation.

   I assign the highest credence to this possibility, because it is hardest to argue against. In contrast to most other theories discussed below, I do not know any way to rule out Lorentzian simplicial quantum gravity (Chapter 2) by comparing with known facts. To know if the extant theories in this category are true, the challenge is to develop techniques to evaluate the path integrals efficiently, draw predictions, gather data, and test against yet unknown facts.

2. Gauge-variable path integrals

   Credence: $\sim 20\%$.

   General relativity can be reformulated in terms of frame fields. This leads to alternative gravitational path integrals based on gauge variables. Examples include Ponzano-Regge models, and various spin foam models.

   For geometric variables, the path integral sum is constrained by inequalities to Lorentzian or Euclidean geometries. For gauge variables, it is far less clear which configurations should be included in the path integral. The many different answers give rise to a large variety of gauge-variable path integrals, of which only a small portion have been studied in detail. The pool is large enough for one to be optimistic that at least one member could be true, but due to limited understanding of the models, there lacks sufficient reason to single out any candidate as particularly promising. Therefore I assign a modest credence.

3. Causal set path integrals

   Credence: $\sim 3\%$.

   The causal set approach captures gravity by fundamentally discrete sets, and provides examples for gravitational path integrals based on neither geometric nor gauge variables.

Causal set path integrals face some long-standing issues. A causal set does not carry a spacetime dimension, nor a tangent space structure to couple to fermions. Consequently, no extant causal set path integral can be viewed as a viable candidate for the true theory of quantum gravity, and the faint hope has to be set on future inventions.

4. Holography

   Credence: $\sim 5\%$.

   As far as our universe goes, AdS/CFT correspondence seems to only supply false theories, because our universe is not AdS.

   There is the hope that some dS holography theory applicable to our universe will be invented in the future. So far so bad, after many years, so my credence is low.

5. Wheeler-DeWitt equation

   Credence: $\sim 1\%$.

   Defining a theory of quantum gravity (e.g., Canonical loop quantum gravity; quantum geometrodynamics) by the Wheeler-DeWitt equation $H\psi = 0$ seems to me to be misguided from the beginning, due to its focus on ψ, instead of empirical predictions. Unlike in undergraduate quantum mechanics, where $p = |\psi|^2$, there is no clear way to reach empirical probabilistic predictions p from ψ in the present context.

   To derive empirical probabilities, one could adopt a functional integral approach (Chapter 6). If one insists on a differential equation approach, one could parameterize the family of mathematical quantities which do yield empirical probabilities, and try to derive an equation of how these quantities change as the parameters vary. This equation may have little to do with the Wheeler-DeWitt equation, which means one should not focus on the Wheeler-DeWitt equation in the first place.

6. Other known approaches

   Credence: $\sim 5\%$.

   The other extant approaches I am know of are all quite premature. For instance, functional renormalization group asymptotic safety, developed in the Euclidean setting, is inapplicable in the Lorentzian setting. Perturbative approaches, limited to the perturbative setting, are capable of neither offering a full theory of quantum gravity, nor addressing questions about black hole interior and quantum cosmology, where quantum gravity is actually in need. String theory turns to holography or the distant dream of M-theory to address the challenge for a non-perturbative formulation. The former case does not seem promising, as discussed above, while the latter dream seems to remain distant.

   Since these are all long-standing issues, unsolved not due to the lack of trying, my credence is low.

7. Something else

   Credence: ?

   It could be that the true theory is not a quantum theory, or a quantum theory in some brand new approach. I find it difficult to estimate the chance for this possibility, so leave it unspecified.

1.2.3 Boundary condition

The task of determining the boundary condition of the universe is mostly actively studied in quantum cosmology. I will discuss the possibilities in detail in Chapter 7. Very briefly, my guess for the possibilities are:

1. No-boundary proposal

   Credence: $< 1\%$.

   Issues: Euclidean path integral diverges; no first-principle justification for complex contours; ambiguity in the choice of contours.

2. Tunnelling proposal

   Credence: $< 10\%$.

   Issues: Ambiguity beyond simple models; lack of first-principle justification.

3. Indifference proposal (Chapter 7)

   Credence: $< 20\%$.

   Issues: Ambiguities in the boundary condition; How does time asymmetry arise?

4. New no-boundary proposal

   Credence: $< 1\%$.

   Issues: The Wheeler-DeWitt equation framework is misguided to start with (Section 1.2.2).

5. Other proposals for quasi-Lorentzian path integrals

   Credence: $< 10\%$.

   Issues: Ambiguity beyond simple models; lack of first-principle justification.

6. Something else

   Credence: ?

1.2.4 Experience

The task of relating theory to experience is mostly actively considered in quantum foundations. Specifically, candidate interpretations can usually be recast as possibilities to relate theory to experience. I will discuss the possibilities in detail in Chapter 6. Very briefly, my guess for the possibilities are:

1. Decoherent histories

   Credence: $< 3\%$.

   Issues: How to select histories to embed experiences? Why embed experiences in histories at all?

2. Everettian interpretations

   Credence: $< 3\%$.

   Issues: How to justify the Schrödinger equation in quantum gravity? How does time-oriented branching structure arise from no time-orientation? Are the reasons given for the emergence of probability in a deterministic setting really valid?

3. Copenhagenish interpretations (including QBism and relational quantum mechanics)

   Credence: $\sim 5\%$.

   Issues: How to phrase the interpretations in terms of the physical theories (e.g., Standard Model) at all? In terms of the physical theories, what are agents, interaction, ψ, observer, and/or fact? Once formulated in terms of the physical theories, what non-superfluous element do the interpretations add at all?

4. Minimal prescription (Chapter 6)

   Credence: $\sim 40\%$.

   Issues: Can the theory of experience invoked to determine selection rules ever be found?

5. Collapse models and Bohmian mechanics

   Credence: $< 5\%$.

   Issues: extant model are already falsified by empirical data from particle physics.

6. Something else

   Credence: $\sim 50\%$.

1.3 The big picture

Combining the most likely cases from the above assessment yields the following picture for a theory of everything.

The dynamical laws for matter and gravity are captured by a joint matter-gravity path integral (Chapter 2, Chapter 6). The boundary condition is such that all configurations are summed over indifferently (Chapter 7). The interior condition selects for individual experiences and nothing else (Chapter 6).

A human life is characterized by a sequence of experiences. Probabilistic predictions for each individual experience can be found in path integrals of the above form, but no collective account for joint experiences can be given (Chapter 6). Individual experiences are related by their enabling conditions, which encode memories (Chapter 6). Comprehension of objects are constructions, derived from regularities and repetitive patterns among the experiences.

1.4 Outline and references

Here is a broad outline of the rest of the thesis, which exposes the above big picture in more detail:

- Chapter 2 based on [2] introduces Lorentzian simplicial quantum gravity.
- Chapter 3 based on [3] and Chapter 4 based on [4] present some modest attempts toward meeting the challenge mentioned in Section 1.2.2, namely, to develop techniques to evaluate the path integral efficiently, draw predictions, gather data, and test against yet unknown facts.

  Some additional relevant works (left out of the thesis for brevity) can be found in [5, 6]. In [7], it is shown that Lorentzian simplicial quantum gravity realizes superposition of time order.
- Chapter 5 based on [8] explains how singularities are avoided in Lorentzian simplicial quantum gravity. The idea has wider implications, since it applies to all gravitational path integrals that exclude singular spacetimes by definition.
- Chapter 6 based on [9] presents a prescription to extract empirical predictions from path integral theories of everything. There one could also find criticisms against several popular quantum interpretations in the context of theory of everything.

  Some additional relevant works (left out of the thesis for brevity) can be found in [10, 11].
- Chapter 7 based on a paper to appear develops a proposal for the boundary condition of the universe. The idea is to sum over all path integral configurations indifferently. An application to de Sitter minisuperspace model illustrates how the expected empirical predictions could be obtained.
- Chapter 8 based on [1] presents a proposal to understand the Standard Model for particle physics as a theory of particles and strings, as opposed to fields, in exchange for a better conceptual understanding of gauge theories.

  An additional relevant work (left out of the thesis for brevity) is [12].

Part I

Gravity

Chapter 2

Complex, Lorentzian, and Euclidean simplicial quantum gravity: numerical methods and physical prospects

Evaluating gravitational path integrals in the Lorentzian has been a long-standing challenge due to the numerical sign problem. We show that this challenge can be overcome in simplicial quantum gravity. By deforming the integration contour into the complex, the sign fluctuations can be suppressed, for instance using the holomorphic gradient flow algorithm. Working through simple models, we show that this algorithm enables efficient Monte Carlo simulations for Lorentzian simplicial quantum gravity.

In order to allow complex deformations of the integration contour, we provide a manifestly holomorphic formula for Lorentzian simplicial gravity. This leads to a complex version of simplicial gravity that generalizes the Euclidean and Lorentzian cases. Outside the context of numerical computation, complex simplicial gravity is also relevant to studies of singularity resolving processes with complex semi-classical solutions. Along the way, we prove a complex version of the Gauss-Bonnet theorem, which may be of independent interest.

2.1 Introduction

To define a path integral, one needs to specify a way to enumerate the configurations to be summed over. For a non-relativistic particle, it is common is to introduce a lattice of discrete time steps, sum over piecewise linear paths across these steps, and take the continuum limit of lattice spaces going to zero [13].

For gravity, one could similarly introduce a simplicial lattice, sum over piecewise flat geometries on the lattice characterized by the edge lengths, and take the limit of lattice refinement (Fig. 2.1). Historically, this method follows from Regge's insight [14] to use piecewise flat geometries to approximate curved space(times) at the classiccal level. Regge's classical approach is usually referred to as Regge calculus, or simplicial gravity, while the quantum path integral based on it is usually referred to as quantum Regge calculus, or simplicial quantum gravity [15, 16, 17, 18, 19].

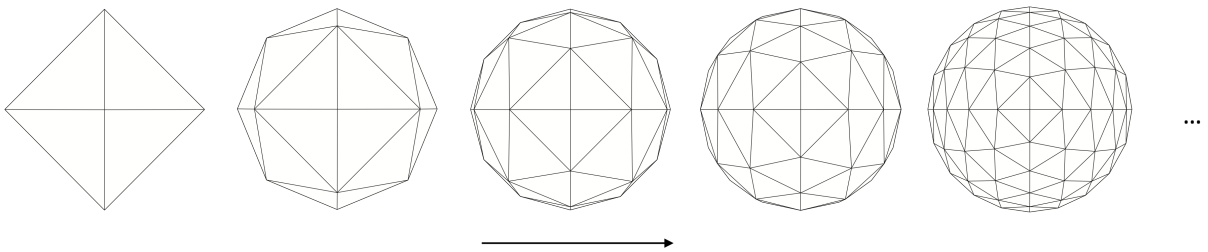

Figure 2.1: Simplicial lattice refinement.

As a non-perturbative path integral approach, simplicial quantum gravity has a clear merit. It is known how to couple to the matter species of the Standard Model (see e.g., Chapter 6 of Hamber's textbook [18] and references therein).

On the other hand, Euclidean quantum gravity faces the conformal instability problem [20]. This is manifested as the problem of the spikes for Euclidean simplicial quantum gravity. In concrete $2D$ models, it is shown that configurations with diverging edge lengths dominate the path integral, even when the total spacetime area is bounded [21]. One view is that only the weak coupling phase is rendered ill by the spiky configurations, but the strong coupling phase stays healthy [22]. A more pessimistic view is that conformal instability poses a lethal threat to Euclidean simplicial quantum gravity.

Whatever conformal instability actually implies about Euclidean quantum gravity, the case is different for the Lorentzian. For $2D$ simplicial quantum gravity it can be shown that the Lorentzian and Euclidean theories are inequivalent, and that spikes are absent in the Lorentzian where spacetime configurations are equipped with causal structures [23, 5]. [1] The question about higher dimensions is open, but the prospect that spikes are absent in the Lorentzian in general, and the fact that spacetime is Lorentzian in Nature form motivations to study Lorentzian simplicial quantum gravity.

Apart from a few works [23, 24, 25, 26, 27, 5], the path integrals of Lorentzian simplicial quantum gravity have not been studied much in the past.[2] Because of the numerical sign problem, naive Monte Carlo simulations do not work efficiently in the Lorentzian as in the Euclidean. This has remained a major obstacle for quantitative studies of Lorentzian simplicial quantum gravity.

In this work we propose to generalize simplicial quantum gravity to the complex domain. This allows us to apply the techniques of complex contour deformation developed in recent years to alleviate the sign problem [30, 31]. By a higher dimensional version of Cauchy's integration theorem, a path integral with a real integration contour can equally be evaluated along a complex contour if the two contours are related across a region where the integrand is holomorphic. The sign problem could be milder on the deformed contour. As reviewed in [31], this idea has been successfully applied to various lattice field theories of matter. It has also been applied to analyze gravitational propagators for spin-foam models in the large spin limit [32].

[1] The proof of the absence of spikes in [23] assumes that the causal signature of simplicial lattice edges are fixed under the path integral. In [5] this assumption is dropped. It is shown that spikes are still absent, provided that causally irregular points with no lightcones attached are prohibited.

[2] In this statement we mean by simplicial quantum gravity the formalism with dynamical lengths. The variation of simplicial quantum gravity with fixed lengths but dynamical lattice graphs has been extensively studied in the form of causal dynamical triangulation [28, 29].

Here we show that the complex contour deformation method also works for Lorentzian simplicial quantum gravity. Monte Carlo simulations are performed to compute the expectation value of spacetime lengths in $1+1D$ using the holomorphic gradient flow algorithm (also called the generalized thimble algorithm) [33, 34, 31]. It is found that the sign fluctuations are largely suppressed on suitable complex contours. As far as we know, this constitutes the first non-perturbative computation of Lorentzian simplicial gravitational path integrals. It opens the possibility to investigate questions about quantum gravity non-perturbatively and quantitatively using Lorentzian simplicial quantum gravity.

Notably, the expectation values computed on the complex contours are directly the results of interest. There is no analytic continuation to Euclidean spacetime like in causal dynamical triangulation [28], nor analytic continuation of parameters in the action like in causal sets [35]. These procedures face the open problem of inverse analytic continuation, which does not arise in the method used here.

Besides overcoming the sign problem, another reason to consider complex simplicial quantum gravity is to study singularity resolving processes. Quantum theory assigns non-zero probabilities to certain processes characterized by boundary conditions admitting not real, but complex semi-classical solutions. A standard example is particle tunneling [36, 37, 38]. It is conceivable that cosmological and black hole singularity resolving processes (see e.g., [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55]) fall into the same category [56, 57, 58, 59, 60, 61, 62, 27]. Lorentzian simplicial quantum gravity provides a formalism to compute the probabilities for such processes. To analyze the semi-classical solutions, the formalism needs to be generalized to the complex domain.

Although simplicial quantum gravity in the complex domain has been studied before [56, 63, 64, 65, 66, 67, 68, 69, 70], the complex theory is reached by analytically continuing the Euclidean theory. In addition, these works concentrated on symmetry-reduced models.

In this work we specify Lorentzian simplicial gravity in arbitrary dimensions and without symmetry reduction with manifestly holomorphic expressions. Upon analytic continuation, the holomorphic expressions define simplicial gravity in the complex domain. The path integrals based on this complex action encompass both Lorentzian and Euclidean simplicial quantum gravity as special cases with different integration contours.

Along the way, we show that the celebrated Gauss-Bonnet theorem admits a complex generalization. This mathematical results may be of independent interest.

The paper is organized as follows. In Section 2.2 and Section 2.3, we review the geometric quantities of length, volume, and areas of Euclidean simplicial gravity, and generalize the quantities to the Lorentzian and complex domains. In Section 2.4 we define simplicial gravitational path integrals in the Lorentzian and complex domains in terms of manifestly holomorphic expressions. In Section 2.5 we review the holomorphic gradient flow algorithm for numerical computations of path integrals with complex actions. Starting in Section 2.6 we specialize to $2D$ simplicial quantum gravity and present the formulas needed for applying the holomorphic gradient flow algorithm. Along the way we prove a complex version of the Gauss-Bonnet theorem. In Section 2.7 we present numerical results that overcome the sign problem. In Section 2.8 we finish with a discussion.

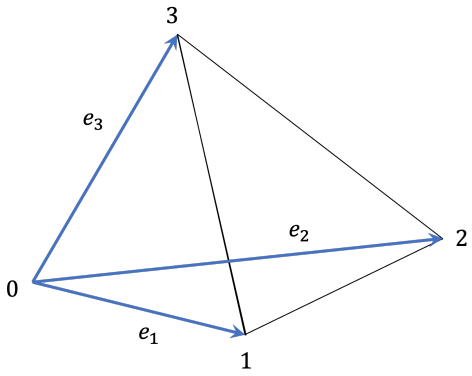

Figure 2.2: A simplex with labelled vertices i and edge vectors e_i.

2.2 Lengths and volumes

In simplicial gravity, the basic variable is the squared length, and the Einstein-Hilbert action is written in terms of volume and angles. (See Hamber's textbook [18] for a comprehensive and lucid introduction to Euclidean simplicial quantum gravity.) In this section and next, we start by presenting length, volume and angles for simplicial geometry in the Euclidean domain, and then generalize these quantities to the Lorentzian and complex domains.

2.2.1 Squared length as the basic variable

Given a metric field g_{ab} on a manifold, the **squared length** σ of a line γ segment is given by

$$\sigma = \int_\gamma ds^2, \tag{2.1}$$

where $ds^2 = g_{ab}dx^a dx^b$ is the line element.

In simplicial gravity each lattice edge e has a squared lengths σ_e with γ taken along the edge. In the Euclidean domain, $\sigma \geq 0$. In the Lorentzian domain, we choose the signature convention that $\sigma > 0$ for spacelike intervals, $\sigma < 0$ for timelike intervals, and $\sigma = 0$ for lightlike intervals.

In a continuum field theory, the basic gravitational variable is usually taken to be the metric field g_{ab}, and the squared length is derived from g_{ab} using (2.1). In contrast, in simplicial gravity the basic variable is usually taken to be the squared lengths σ on the lattice edges. A gravitational configuration is given in terms of the squared length on the edges, from which the metric can be derived as follows.

Let a d-simplex be given and label the vertices by $0, 1, \cdots, d$ (Fig. 2.2). Within the simplex we set up a coordinate system whose basis vectors e_i for $i = 1, \cdots, d$ point from vertex 0 to vertex i. Define a dot product $\cdot$ by

$$e_i \cdot e_j = \frac{1}{2}(\sigma_{0i} + \sigma_{0j} - \sigma_{ij}), \tag{2.2}$$

where σ_{ij} for $i, j = 0, 1, \cdots, d$ are the squared lengths of the edges connecting vertices i and j. Using the metric

$$g_{ij} = \frac{1}{2}(\sigma_{0i} + \sigma_{0j} - \sigma_{ij}), \tag{2.3}$$

the dot product of any pair of vectors $u = u^i e_i$ and $v = v^i e_i$ can be computed as $u \cdot v = g_{ij} u^i v^j$, where the Einstein summation convention is used.

The metric (2.3) is the simplicial analog of the continuum metric. In the continuum, squared lengths are computed through $ds^2 = g_{ab} dx^a dx^b$. On a simplicial lattice, edge squared lengths are computed through

$$\sigma = v \cdot v = g_{ij} v^i v^j, \tag{2.4}$$

where v is the edge vector. For edges containing vertex 0, $v = e_i$, and $v \cdot v = g_{ii} = \sigma_{0i}$. For other edges, $v = e_i - e_j$, and $v \cdot v = g_{ii} - g_{ij} - g_{ji} + g_{jj} = \sigma_{ij}$.

The simplex is understood to have a homogeneous interior. For a line segment within the simplex, the square length is computed by the same formula (2.4) where v is the vector for the line segment.

2.2.2 Complexifying strategy

In complexifying simplicial geometry, we adopt a "squared length based" methodology. After identifying a quantity of interest, such as volumes and angles, we express it as a function of the squared lengths. The function is chosen to agree with known expressions in the Lorentzian and/or Euclidean domains,where the squared lengths take real values. In addition, the function should be holomorphic if possible to facilitate the deformations of integration contours when we study of the quantum theory.

Suppose the above two requirements can be met. Then we can analytically continue the domain of the function to complex squared lengths. When multi-valued functions such as the square root and the log are present, we will extend the domain to be the corresponding Riemann surfaces.

As an example, consider the (linear) **length** defined by $l = \sqrt{\sigma}$. This function is holomorphic away from the branch point $\sigma = 0$. In the Euclidean domain $l > 0$. In the Lorentzian domain $l > 0$ for spacelike edges, and l is positive imaginary for timelike edges in the current choice of the positive branch for the square root.

2.2.3 Volumes

The squared length and length given above are special cases of squared volumes and volumes.

In the continuum, let s be a simplex defined by some unit vectors. Suppose the metric is constant in the region of the simplex. Then the squared volume for the simplex is $\mathbb{V} = \int_s \det g_{ab}(x)\ d^D x = \frac{1}{d!} \det g_{ab}$, where $\frac{1}{d!}$ arises because this is for a simplex rather than a hypercube.

On a simplicial complex, define the **squared volume** of a d-simplex by

$$\mathbb{V} = \frac{1}{(d!)^2} \det g_{ij}, \tag{2.5}$$

where g_{ij} as defined in (2.3) is a function of the edge squared lengths. An equivalent expression that is manifestly symmetric in the squared lengths is the Cayley-Menger determinant

$$\mathbb{V} = \frac{(-1)^{d+1}}{2^d (d!)^2} \begin{vmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & \sigma_{01} & \sigma_{02} & \dots & \sigma_{0d} \\ 1 & \sigma_{01} & 0 & \sigma_{12} & \dots & \sigma_{1d} \\ 1 & \sigma_{02} & \sigma_{12} & 0 & \dots & \sigma_{2d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_{0d} & \sigma_{1d} & \sigma_{2d} & \dots & 0 \end{vmatrix}. \tag{2.6}$$

The **volume** V of a d-simplex is defined by

$$V = \sqrt{\mathbb{V}}. \tag{2.7}$$

Both $\mathbb{V}$ and V are defined for complex squared lengths. In (2.7) the squared volume is taken to live on the Riemann surface of the square root function. V is holomorphic as a function of the squared lengths away from the branch points where $\mathbb{V} = 0$.

In the Euclidean domain, $\mathbb{V} > 0$. In the Lorentzian domain, $\mathbb{V} \leq 0$. The positive branch for the square root is chosen so that V is positive imaginary or zero for Lorentzian simplices.

Example 1. *In lower dimensions some familiar expressions are recovered. In $1D$ the volumes derived from (2.6) and (2.7) are*

$$\mathbb{V} = \sigma_{01}, \tag{2.8}$$
$$V = \sqrt{\sigma_{01}}, \tag{2.9}$$

which reproduce the length formulas. In $2D$ the volumes for a triangle t derived from (2.6) and (2.7) are

$$\mathbb{V} = \frac{1}{16}\left(-\sigma_{01}^2 - \sigma_{02}^2 - \sigma_{12}^2 + 2\sigma_{01}\sigma_{02} + 2\sigma_{01}\sigma_{12} + 2\sigma_{02}\sigma_{12}\right), \tag{2.10}$$
$$V = \frac{1}{4}\sqrt{-\sigma_{01}^2 - \sigma_{02}^2 - \sigma_{12}^2 + 2\sigma_{01}\sigma_{02} + 2\sigma_{01}\sigma_{12} + 2\sigma_{02}\sigma_{12}}, \tag{2.11}$$

which reproduce Heron's formula for triangle areas. □

2.2.4 Generalized triangle inequalities

The squared distances must obey certain generalized triangle inequalities to describe Euclidean and Lorentzian simplices.

In the Euclidean domain, a simplex s obeys

$$\mathbb{V} > 0 \quad \text{for all subsimplices of } s \text{ including } s \text{ itself.} \tag{2.12}$$

For example, for a triangle this means the squared area and the squared lengths are positive:

$$\mathbb{V} = \frac{1}{16}\left(-\sigma_{01}^2 - \sigma_{02}^2 - \sigma_{12}^2 + 2\sigma_{01}\sigma_{02} + 2\sigma_{01}\sigma_{12} + 2\sigma_{02}\sigma_{12}\right) > 0, \tag{2.13}$$
$$\sigma_{01}, \sigma_{02}, \sigma_{03} > 0. \tag{2.14}$$

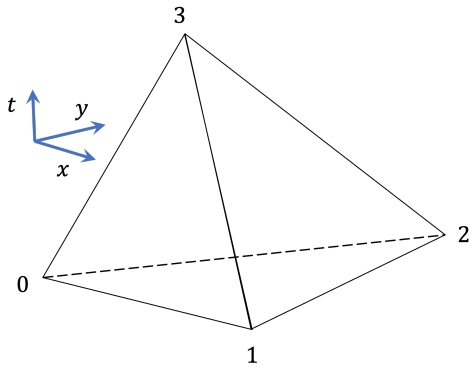

Figure 2.3: A $3D$ timelike simplex can have a spacelike subsimplex 012 in addition to timelike subsimplices such as 013.

In the Lorentzian domain, a simplex s obeys [24, 26]

$$\mathbb{V}_s < 0; \text{ and } \mathbb{V}_r < 0 \implies \mathbb{V}_t <= 0 \text{ for all } t \supset r. \tag{2.15}$$

A simplex is timelike if $\mathbb{V} < 0$, and spacelike if $\mathbb{V} > 0$. In contrast to the Euclidean domain where all simplices and subsimplices have the same causal signature (spacelike), in the Lorentzian domain the subsimplices are allowed to be both timelike and spacelike Figure 2.3. The Lorentzian generalized inequalities (2.15) first say that the simplex s itself needs to be timelike. Furthermore, if any subsimplex r is timelike, then all subsimplices t containing r cannot be spacelike. This is because a timelike subsimplex cannot be embedded in a spacelike subsimplex. For instance in Figure 2.3, if the edge subsimplex 03 is timelike, then the triangle subsimplices 013 and 023 containing the timelike edge 03 must not be spacelike, which is a reasonable condition.

2.3 Angles

2.3.1 Euclidean angles

In Euclidean space, what is the angle θ bounded by two vectors a and b? Since

$$a \cdot b = |a||b| \cos\theta, \quad |x| := \sqrt{x \cdot x}, \tag{2.16}$$

one answer is that $\theta = \cos^{-1} \frac{a \cdot b}{|a||b|}$. Another answer is in terms of the scalar wedge product defined by

$$a \wedge b = \sqrt{(a \cdot b)^2 - (a \cdot a)(b \cdot b)}. \tag{2.17}$$

Using $\sin^2\theta + \cos^2\theta = 1$, it is easy to see that for $\theta > 0$,

$$a \wedge b = i|a||b| \sin\theta. \tag{2.18}$$

Therefore $\theta = \sin^{-1} \frac{a \wedge b}{i|a||b|}$.

The answer (2.16) or (2.18) in isolation has ambiguities, because different angles can have the same $\cos$ or $\sin$ values. Within a 2π period, angles are uniquely determined when the information of $\cos^{-1}$ and $\sin^{-1}$ are combined. From (2.16) and (2.18), we derive that $e^{i\theta} = \frac{1}{|a||b|}(a \cdot b + a \wedge b)$, so[3]

$$\theta = -i \log \alpha, \tag{2.19}$$

$$\alpha = \frac{1}{|a||b|}(a \cdot b + a \wedge b). \tag{2.20}$$

This determines θ uniquely within a 2π period depending on the choice of the branch for the log function.

2.3.2 Complex angles

In the general complex domain, we take

$$\theta = -i \log \alpha, \tag{2.21}$$

$$\alpha = \frac{a \cdot b + a \wedge b}{\sqrt{a \cdot a}\sqrt{b \cdot b}} = \frac{a \cdot b + \sqrt{(a \cdot b)^2 - (a \cdot a)(b \cdot b)}}{\sqrt{a \cdot a}\sqrt{b \cdot b}}, \tag{2.22}$$

as the definition of **complex angles**. Equation (2.22) is one of the expressions in Sorkin's definition of Lorentzian angles in the Minkowski plane [71].[4] Here we recognize that more generally, (2.21) and (2.22) offer a unified definition for Euclidean, Lorentzian, and complex angles in all cases.[5]

In terms of the edge squared lengths (Fig. 2.4),

$$a \cdot b = \frac{1}{2}(\sigma_a + \sigma_b - \sigma_c), \tag{2.23}$$

$$a \cdot a = \sigma_a, \quad b \cdot b = \sigma_b, \tag{2.24}$$

$$a \wedge b = \frac{1}{2}\sqrt{\sigma_a^2 + \sigma_b^2 + \sigma_c^2 - 2\sigma_a\sigma_b - 2\sigma_b\sigma_c - 2\sigma_c\sigma_a}. \tag{2.25}$$

Therefore

$$\theta = -i \log \alpha,$$

$$\alpha = \frac{\sigma_a + \sigma_b - \sigma_c + \sqrt{\sigma_a^2 + \sigma_b^2 + \sigma_c^2 - 2\sigma_a\sigma_b - 2\sigma_b\sigma_c - 2\sigma_c\sigma_a}}{2\sqrt{\sigma_a}\sqrt{\sigma_b}}. \tag{2.26}$$

We take (2.26) as the definition of **complex angles** in terms of complex squared lengths. This function is holomorphic away from the log and square root branch points. At the square root branch point of $a = 0$ or $b = 0$, the denominator becomes 0. We will comment more on the (ir)relevance of this case in the end of Section 2.3.3 and in Section 2.7.3.

[3] This formula is related to the so-called "geometric product" $\vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b}$, which offers a way to encode rotations. The difference is that here $\vec{a} \wedge \vec{b}$ is a bivector instead of a scalar.

[4] In Sorkin's definition of Lorentzian triangles [71], (2.22) is used for angles bounded by two spacelike vectors in the same quadrant, and angles bounded by a spacelike vector and a timelike vector. A different expression is used for angles bounded by two timelike vectors in the same quadrant.

[5] For the formula to apply to the Euclidean case, the $-i$ factor in (2.21) is necessary. In comparing with other works based on Sorkin's definition one should keep in mind that the $-i$ factor is absent there. In addition, for Lorentzian angles (2.28) defined below differs in the choice of square root branches from Sorkin's formula.

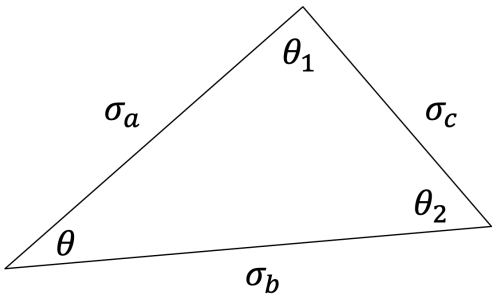

Figure 2.4: A triangle with squared lengths $\sigma_a, \sigma_b, \sigma_c$.

Note from (2.10) that the input A to the numerator square root equals $-16\mathbb{V}$, where $\mathbb{V}$ is the squared volume for the triangle in Fig. 2.4. By the triangle inequalities of Section 2.2.4, $A > 0$ for a Lorentzian triangle and $A < 0$ for an Euclidean triangle.

For Euclidean angles the principal branches of the log and square root functions are chosen. The complex angles then reduce to the correct Euclidean angles, since the former are obtained by generalizing the latter. The choices of branches for Lorentzian angles are specified below.

Sum of complex angles in a triangle

The angles of an Euclidean triangle sum to π. In the complex domain, this generalizes to $(2n+1)\pi$ with $n \in \mathbb{Z}$.

Proposition 2. *The complex angles sum to $(2n+1)\pi$ with $n \in \mathbb{Z}$ for a triangle of complex squared edge lengths.*

Proof. Consider a triangle with complex squared lengths $\sigma_a, \sigma_b, \sigma_c$ (Figure 2.4), and complex angles $\theta = -i \log \alpha, \theta_1 = -i \log \alpha_1, \theta_2 = -i \log \alpha_2$. A straightforward calculation using (2.26) yields

$$\alpha_1\alpha_2 = \frac{-\sigma_a - \sigma_b + \sigma_c + \sqrt{\sigma_a^2 + \sigma_b^2 + \sigma_c^2 - 2\sigma_a\sigma_b - 2\sigma_b\sigma_c - 2\sigma_c\sigma_a}}{2\sqrt{\sigma_a}\sqrt{\sigma_b}}. \tag{2.27}$$

A similar calculation yields $\alpha\alpha_1\alpha_2 = -1$. For the complex log function, $\log(z_1 z_2) = \log z_1 + \log z_2$ up to multiples of $2\pi i$. Therefore $\theta + \theta_1 + \theta_2 = -i \log(-1) + 2\pi n = (2n+1)\pi$, where n is an integer. □

2.3.3 Lorentzian angles

In this section, we consider angles for Lorentzian simplicial geometries that obey the Lorentzian generalized triangle inequalities (2.15). In previous works [71, 23, 26], not one, but multiple expressions for Lorentzian angles in terms of log and trigonometric functions were used depending on where the edges lie in the Minkowski plane. A merit of the complex angle defined above is that it unifies these multiple cases (as well as the Euclidean case) in one formula.

Here we focus on convex angles, because in simplicial gravity only these arise from individual simplices. Non-convex angles arise from summing the convex angles of individual simplices. Here we consider the branch choice

$$\begin{aligned}\theta =& -i \operatorname{Log} \alpha, \\ \alpha =& \frac{a\cdot b + \sqrt{(a\cdot b)^2 - (a\cdot a)(b\cdot b)}}{\sqrt{a\cdot a - 0i}\sqrt{b\cdot b - 0i}},\end{aligned} \tag{2.28}$$

for Lorentzian angles. In terms of the squared lengths,

$$\alpha = \frac{\sigma_a + \sigma_b - \sigma_c + \sqrt{\sigma_a^2 + \sigma_b^2 + \sigma_c^2 - 2\sigma_a\sigma_b - 2\sigma_b\sigma_c - 2\sigma_c\sigma_a}}{2\sqrt{\sigma_a - 0i}\sqrt{\sigma_b - 0i}}. \tag{2.29}$$

Here

$$\operatorname{Log} z = \log r + i\phi, \quad z = re^{i\phi} \text{ with } \phi \in (-\pi, \pi] \tag{2.30}$$

$$\sqrt{z} = \sqrt{r}e^{i\phi/2}, \quad z = re^{i\phi} \text{ with } \phi \in (-\pi, \pi], \tag{2.31}$$

$$\sqrt{z - 0i} = \sqrt{r}e^{i\phi/2}, \quad z = re^{i\phi} \text{ with } \phi \in [-\pi, \pi). \tag{2.32}$$

The first two are just the principal branches of log and square root. The third one $\sqrt{z-0i}$ is negative imaginary for $z < 0$. The symbol $-0i$ is a reminder that $z < 0$ is continuously connected to $z > 0$ through the lower complex plane instead of the upper one.

The following properties hold for Lorentzian angles.

Proposition 3. *The Lorentzian convex angles θ defined by formula (2.28) are additive.*

Proposition 4. *The complex Lorentzian angle θ is related to the Lorentz boost angle θ_{boost} by*

$$\theta = -i\theta_{boost}. \tag{2.33}$$

Here the convention is that $\theta_{boost} > 0$ for a boost angle relating spacelike vectors, and $\theta_{boost} < 0$ for a boost angle relating timelike vectors.

Proposition 5. *Between two edges related by the reflection across a light ray, the angle θ equals*

$$\theta = \pi/2, \tag{2.34}$$

whose imaginary part vanishes.

Proposition 6. *In the flat Minkowski plane, the angles around a point sum to 2π.*

Proposition 7. *For a convex Lorentzian angle θ,*

$$\operatorname{Re}\theta = N\pi/2, \tag{2.35}$$

where $N = 0, 1, 2$ is the number of light rays enclosed within the angle.

Proposition 8. *The angles of a Lorentzian triangle sum to 2π.*

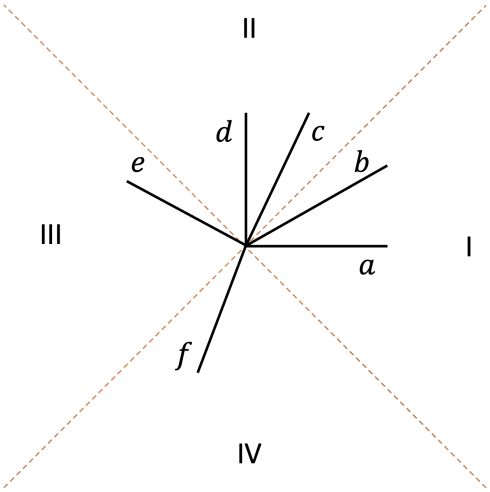

Figure 2.5: The Minkowski plane with four quadrants bounded by dashed light rays. The edges a to f are distributed in different quadrants.

These results are easier to derive after working through some examples. These also serve to help readers unfamiliar with Lorentzian angles [71] to build some intuitions.

In the Minkowski plane, a convex angle can bound $N = 0, 1,$ or 2 light rays (Fig. 2.5). According to whether the vectors bounding the angle are timelike or spacelike (for reasons mentioned below all the examples, we do not consider lightlike edges here), there are five cases in total. We consider them in turn.

Example 9 (Spacelike edges within the same quadrant)**.** *Consider spacelike edges a and b forming a triangle with squared lengths $\sigma_a = 1, \sigma_b = 3/4, \sigma_{ab} = -1/4$, where σ_{ab} is the squared length for the third edge (Figure 2.5). The complex angle θ bounded by a and b can be calculated using (2.23) to (2.25) as follows.*

$$a \cdot b = \frac{1}{2}(\sigma_a + \sigma_b - \sigma_{ab}) = 1, \tag{2.36}$$

$$a \cdot a = \sigma_a = 1, \quad b \cdot b = \sigma_b = 3/4, \tag{2.37}$$

$$a \wedge b = \sqrt{(a \cdot b)^2 - (a \cdot a)(b \cdot b)} = 1/2, \tag{2.38}$$

$$\begin{aligned} \theta &= -i \log\left(\frac{a \cdot b + a \wedge b}{\sqrt{a \cdot a - 0i}\sqrt{b \cdot b - 0i}} \right) \\ &= -i \log \sqrt{3}. \end{aligned} \tag{2.39}$$

□

The above calculation is based on the invariant quantity of the squared length and does not invoke any coordinate system. Alternatively, one could introduce a coordinate system in the Minkowski plane, represent a and b as vectors there, and use the Minkowski inner product for $\cdot$ to calculate θ. For instance, one could choose $a = (1, 0)$ and $b = (1, 1/2)$ in the coordinate convention (x, t). Then again $a \cdot b = 1^2 - 0 = 1$, $a \cdot a = 1^2 - 0^2 = 1$, and $b \cdot b = 1^2 - (1/2)^2 = 3/4$, so one will get the same result for θ.

The complex angle θ is related to the boost angle of Lorentz transformations. The boost angle from a to b is, up to a choice of sign,

$$\theta_{\text{boost}} = \cosh^{-1}(\hat{a} \cdot \hat{b}). \tag{2.40}$$

Here $\hat{x} := x/\sqrt{|x \cdot x|}$ denotes the normalized vector for x. For the vectors $a = (1, 0)$ and $b = (1, 1/2)$, $\hat{a} = (1, 0)$ and $\hat{b} = (2/\sqrt{3}, 1/\sqrt{3})$. Therefore $\theta_{\text{boost}} = \cosh^{-1}(2/\sqrt{3}) = \log\sqrt{3}$, where we used the elementary identity

$$\cosh^{-1} z = \log\Big(z + \sqrt{z^2 - 1}\Big). \tag{2.41}$$

In this case for two spacelike edges, upon choosing the boost angle to be positive, we see that $\theta = -i\theta_{\text{boost}}$. Using (2.40) and (2.41), it is easy to check that this relation holds for all pairs of spacelike edges in the same quadrant in the Minkowski plane. We will see next that this relation also holds for timelike edges.

Example 10 (Timelike edges within the same quadrant)**.** *Consider timelike edges c and d forming a triangle with squared lengths $\sigma_c = -3/4, \sigma_d = -1, \sigma_{cd} = 1/4$, where σ_{cd} is the squared length for the third edge (Figure 2.5). The complex angle θ bounded by c and d can be calculated using (2.23) to (2.25) as follows.*

$$\begin{aligned}
c \cdot d =& \frac{1}{2}(\sigma_c + \sigma_d - \sigma_{cd}) = -1, && (2.42)\\
c \cdot c =& \sigma_c = -3/4, \quad d \cdot d = \sigma_b = -1, && (2.43)\\
c \wedge d =& \sqrt{(c \cdot d)^2 - (c \cdot c)(d \cdot d)} = 1/2, && (2.44)\\
\theta =& -i \log\left(\frac{c \cdot d + c \wedge d}{\sqrt{c \cdot c - 0i}\sqrt{d \cdot d - 0i}}\right)\\
=& -i \log \frac{-1 + 1/2}{(-i\sqrt{3/4})(-i\sqrt{1})} = -i \log\Big(1/\sqrt{3}\Big). && (2.45)
\end{aligned}$$

□

Alternatively, setting $c = (1/2, 1)$ and $d = (0, 1)$ in a coordinate system (x, t) and performing the calculation there leads to the same θ.

Note that c and b, as well as d and a are related by reflection with respect to the light ray separating quadrant I and II. The same Lorentz boost transformation that maps a to b will map d to c. The boost angle from a to b is anti-clockwise, while that from d to c is clockwise. Since we chose the boost angle from a to b to be positive, it is reasonable to choose the boost angle from d to c to be negative. In this case we have

$$\theta_{\text{boost}} = -\cosh^{-1}(|\hat{c} \cdot \hat{d}|) = -\cosh^{-1}(-\hat{c} \cdot \hat{d}), \tag{2.46}$$

since for timelike vectors in the same quadrant $\hat{c} \cdot \hat{d} < 0$, and the normalized vectors take the form $\hat{x} := x/\sqrt{|x \cdot x|} = x/\sqrt{-x \cdot x}$. From this we obtain $\hat{c} = (1/\sqrt{3}, 2/\sqrt{3})$ and $\hat{d} = (0, 1)$, so $\theta_{\text{boost}} = -\cosh^{-1}(2/\sqrt{3}) = \log\big(1/\sqrt{3}\big)$.

Again, $\theta = -i\theta_{\text{boost}}$. Using (2.46) and (2.41), it is not hard to check that actually this relation holds for all pairs of timelike edges in the same quadrant in the Minkowski plane. Since boost angles exist only between two spacelike vectors in the same quadrant and two timelike vectors in the same quadrant, we have proved Theorem 4.

Example 11 (A spacelike edge and a time like edge)**.** *Consider the spacelike edge a and timelike edge c forming a triangle with squared lengths $\sigma_a = 1, \sigma_c = -3/4, \sigma_{ac} = -3/4$, where σ_{ac} is the squared length for the third edge (Figure 2.5). The complex angle θ bounded by a and c can be calculated using (2.23) to (2.25) as follows.*

$$a \cdot c = \frac{1}{2}(\sigma_a + \sigma_c - \sigma_{ac}) = 1/2, \tag{2.47}$$

$$a \cdot a = \sigma_a = 1, \quad c \cdot c = \sigma_c = -3/4, \tag{2.48}$$

$$a \wedge c = \sqrt{(a \cdot c)^2 - (a \cdot a)(c \cdot c)} = 1, \tag{2.49}$$

$$\begin{aligned} \theta &= -i \log\left(\frac{a \cdot c + a \wedge c}{\sqrt{a \cdot a - 0i}\sqrt{c \cdot c - 0i}} \right) \\ &= -i \log \frac{1/2 + 1}{(\sqrt{1})(-i\sqrt{3/4})} = -i \log\left(i\sqrt{3}\right) = -i \log \sqrt{3} + \pi/2. \end{aligned} \tag{2.50}$$

□

Alternatively, setting $a = (1, 0)$ and $c = (1/2, 1)$ in a coordinate system (x, t) and performing the calculation there leads to the same θ.

Note the relevance of the choice of branch for the square root. Had we chosen the branch without $-0i$, the denominator would be $i\sqrt{3/4}$ instead, and the real part of θ would be $-\pi/2$. In the choice with $-0i$, we have:

Lemma 12. *The angle θ between a spacelike edge and a timelike edge obeys*

$$\operatorname{Re} \theta = \pi/2. \tag{2.51}$$

Proof. Without loss of generality let $\sigma_a > 0$ and $\sigma_c < 0$. Then $(a \cdot a)(c \cdot c) < 0$, so $a \wedge c = \sqrt{(a \cdot c)^2 - (a \cdot a)(c \cdot c)} > |a \cdot c|$. Therefore the numerator of α, $a \cdot c + a \wedge c$, is positive. The denominator $\sqrt{a \cdot a - 0i}\sqrt{c \cdot c - 0i}$ is negative imaginary. Therefore α is positive imaginary. It follows that $\theta = -i \log(ir) = -i \log r + \pi/2$ for some $r > 0$. □

For the special case of two edges a and c related by a reflection across a light ray as the reflection axis, the angle bounded by them equals $\theta = \pi/2$. This is the content of Theorem 5, which is proved by noting that $\sigma_a = -\sigma_c$ and $\sigma_{ac} = 0$. From these we derive that $a \cdot c = 0$, $a \wedge c = \sqrt{\sigma_a^2}$, whence $\alpha = \sqrt{\sigma_a^2}/(\sqrt{\sigma_a - 0i}\sqrt{-\sigma_a - 0i}) = i$. Therefore $\theta = -i \log \alpha = \pi/2$.

This should be expected. The boost angles from a and c to the light ray are equal in magnitude and opposite in sign. When added up to obtain $\operatorname{Im} \theta$ according to Theorem 4, they cancel. By Theorem 7, $\operatorname{Re} \theta = \pi/2$ because travelling from a to c crosses one light ray.

Theorem 5 implies that the in the flat Minkowski plane the angles around around a point sum to 2π, which is the content of Theorem 6. Consider four edges right in the middle of the four quadrants. According to Theorem 5, the four angles formed by them all equal $\pi/2$, so they sum to 2π.

Example 13 (Spacelike edges in different quadrants)**.** *Consider two spacelike edges a and e in different quadrants forming a triangle with squared lengths $\sigma_a = 1, \sigma_e = 3/4, \sigma_{ae} = 15/4$, where*

σ_{ae} is the squared length for the third edge (Figure 2.5). The complex angle θ bounded by a and e can be calculated using (2.23) to (2.25) as follows.

$$\begin{aligned} a \cdot e =& \frac{1}{2}(\sigma_a + \sigma_e - \sigma_{ae}) = -1, && (2.52)\\ a \cdot a =& \sigma_a = 1, \quad e \cdot e = \sigma_e = 3/4, && (2.53)\\ a \wedge e =& \sqrt{(a \cdot e)^2 - (a \cdot a)(e \cdot e)} = 1/2, && (2.54)\\ \theta =& -i \log\left(\frac{a \cdot e + a \wedge e}{\sqrt{a \cdot a - 0i}\sqrt{e \cdot e - 0i}} \right) \\ =& -i \log\left(-1/\sqrt{3}\right) = -i \log\left(1/\sqrt{3}\right) + \pi. && (2.55) \end{aligned}$$

□

Again, the readers can check that the vectors $a = (1, 0)$ and $e = (-1, 1/2)$ leads to the same θ.

Note the relevance of the choice of branch for the log function. The principal branch which we chose yields $\mathrm{Re}\,\theta = \pi$ for $\alpha < 0$. A different choice could result in $\mathrm{Re}\,\theta = -\pi$. Given the branch choices for the square roots, only for the principal branch can the angles possibly be additive. To see this, note that by Theorem 12, each light ray crossing accrues $\pi/2$ for $\mathrm{Re}\,\theta$. Since from a to e there are two light rays crossed, $\mathrm{Re}\,\theta$ needs to be π if the angles are additive.

Example 14 (Timelike edges in different quadrants)**.** *Consider two timelike edges d and f in different quadrants forming a triangle with squared lengths $\sigma_d = -1, \sigma_f = -3/4, \sigma_{df} = -15/4$, where σ_{df} is the squared length for the third edge (Figure 2.5). The complex angle θ bounded by d and f can be calculated using (2.23) to (2.25) as follows.*

$$\begin{aligned} d \cdot f =& \frac{1}{2}(\sigma_d + \sigma_f - \sigma_{df}) = 1, && (2.56)\\ d \cdot d =& \sigma_d = -1, \quad f \cdot f = \sigma_f = -3/4, && (2.57)\\ d \wedge f =& \sqrt{(d \cdot f)^2 - (d \cdot d)(f \cdot f)} = 1/2, && (2.58)\\ \theta =& -i \log\left(\frac{d \cdot f + d \wedge f}{\sqrt{d \cdot d - 0i}\sqrt{f \cdot f - 0i}} \right) \\ =& -i \log \frac{1 + 1/2}{(-i)(-i\sqrt{3/4})} = -i \log\left(-\sqrt{3}\right) = -i \log\left(\sqrt{3}\right) + \pi. && (2.59) \end{aligned}$$

□

Alternatively, setting $d = (0, 1)$ and $f = (-1/2, -1)$ in a coordinate system (x, t) and performing the calculation there leads to the same θ.

In the above two cases $\mathrm{Re}\,\theta = \pi$. It actually holds in general that crossing two light rays makes the angle accrue a real part of π. The reason is that the log argument is negative for two light ray crossings, which yields $\mathrm{Re}\,\theta = \pi$. To see that the log argument is negative, note that for two spacelike vectors a and e in different quadrants, $a \cdot e = \frac{1}{2}(\sigma_a + \sigma_e - \sigma_{ae}) < 0$ as a consequence of the Lorentzian triangle inequality (2.15). Therefore the log argument $\frac{a \cdot e + a \wedge e}{\sqrt{a \cdot a - 0i}\sqrt{e \cdot e - 0i}} < 0$. For two timelike vectors d and f in different quadrants, $d \cdot f = \frac{1}{2}(\sigma_d + \sigma_f - \sigma_{df}) > 0$ as a consequence of the Lorentzian triangle inequality (2.15). In addition, $d \wedge f = \sqrt{(d \cdot f)^2 (d \cdot d)(f \cdot f)} < |d \cdot f|$. Therefore the log argument $\frac{d \cdot f + d \wedge f}{\sqrt{d \cdot d - 0i}\sqrt{f \cdot f - 0i}} < 0$.

Since a convex angle in the Minkowski plane can only enclose $0, 1$ or 2 light rays, we have proved Theorem 7. For any triangle in the Minkowski plane, the three angles enclose two light rays in total. Therefore the sum of the three angles have π as the real part. By Theorem 2, the imaginary part vanishes. This proves Theorem 8.

Finally, we want to prove Theorem 3, i.e.,

$$\theta(a,c) = \theta(a,b) + \theta(b,c), \tag{2.60}$$

where b lies between a and c in the Minkowski plane, and $\theta(x,y) = -i\log\alpha(x,y)$ denotes the convex angle defined by some vectors x and y according to (2.28). The first part of the proof is the same as Sorkin's proof for his equation (3) in [71]. Explicitly, since the angles are convex and b lies in between a and c, one could write $b = \alpha a + \beta c$ with $\alpha, \beta \geq 0$. This can be plugged in

$$(\frac{a\cdot b + a\wedge b}{\sqrt{a\cdot a}\sqrt{b\cdot b}})(\frac{b\cdot c + b\wedge c}{\sqrt{b\cdot b}\sqrt{c\cdot c}}) = \frac{a\cdot c + a\wedge c}{\sqrt{a\cdot a}\sqrt{c\cdot c}}, \tag{2.61}$$

i.e., $\alpha(a,b)\alpha(b,c) = \alpha(a,c)$, to eliminate b and establish the identity.

For the complex log function, $\theta(a,b)+\theta(b,c) = -i\log\alpha(a,b) - i\log\alpha(b,c) = -i\log(\alpha(a,b)\alpha(b,c)) = -i\log\alpha(a,c) = \theta(a,c)$ up to an integer multiple of 2π. However, by Theorem 7 and the assumption that all three angles are convex, the real part of the left hand side can only be $0, \pi/2$, or π. The same holds for the right hand side. Therefore the multiple of 2π has to be zero, and we established $\theta(a,b) + \theta(b,c) = \theta(a,c)$.

Lightlike edges

When one or two of the edges that bound the angle are lightlike, the Lorentzian angle defined in (2.28) could diverge. In [71], special care is taken to redefine such angles.

We will not perform any redefinition for angles with lightlike edges in this work, because the main focus is on the quantum theory. In the path integral, squared lengths is integrated over for each edge. Zero (lightlike) squared length is of measure zero, and a special redefinition just on this measure zero set is not necessary. See Section 2.7.3 for additional discussions on the (ir)relevance of lightlike edges for the gravitational path integral.

2.3.4 Dihedral angles

In simplicial gravity, curvature is captured by deficit angles, which is in turn defined in terms of dihedral angles.

A dihedral angles is formed by two codimension-1 faces at a hinge, which is a codimension-2 simplex. For instance in $2D$, the dihedral angle $\theta_{s,h}$ in triangle s at vertex h is the angle formed by the two edges sharing h. In $3D$ the dihedral angle $\theta_{s,h}$ in tetrahedron s at edge h is the angle formed by the two triangles sharing h. In $4D$ the dihedral angle $\theta_{s,h}$ in 4-simplex s at triangle h is the angle formed by the two tetrahedrons sharing h etc.

As illustrated in Figure 2.6, dihedral angles can be obtained by projecting s to the triangle orthogonal to h, and extracting the triangle angle at the vertex that h projects to. Using (2.22), namely

$$\theta = -i\log\alpha, \quad \alpha = \frac{a\cdot b + \sqrt{(a\cdot b)^2 - (a\cdot a)(b\cdot b)}}{\sqrt{a\cdot a}\sqrt{b\cdot b}}, \tag{2.62}$$

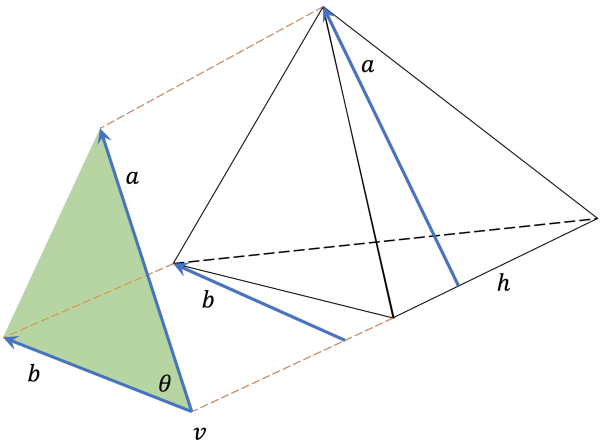

Figure 2.6: In $3D$, the tetrahedron simplex s projects into the shaded triangle orthogonal to the hinge edge h. The dihedral angle $\theta_{s,h}$ projects to the triangle angle θ. The faces bounding the dihedral angle project to the edges a and b of the triangle.

the dihedral angle can be computed from $a \cdot b$, $a \cdot a$, and $b \cdot b$ of the projected triangle. However, in simplicial gravity the input data are the squared distances σ_e on the simplicial edges e. We need to express $a \cdot b$, $a \cdot a$, and $b \cdot b$ in terms σ_e.

Volume forms

To express $a \cdot b$, $a \cdot a$, and $b \cdot b$ in terms σ_e, it is useful to introduce a volume form representation of the (sub)simplices [72]. An n-simplex has $n+1$ vertices. With one of the vertices labelled as 0, the n vectors $e_i, i = 1, \ldots, n$ starting from 0 and pointing to the other n vertices characterize the simplex (Fig. 2.2).

In Section 2.2.3 we treated e_i as the basis vectors in defining the metric g_{ij} which equals $e_i \cdot e_j$. Let e^i be the dual vectors so that $e^i(e_j) = \delta^i_j$. A d-simplex s can represented by the d-form

$$\omega_s = e^1 \wedge \cdots \wedge e^d. \tag{2.63}$$

Then an n-dimensional subsimplex r with edge vectors $e_{r_1}, \ldots, e_{r_n}$ can be represented by the n-form

$$\omega_r = e^{r_1} \wedge \cdots \wedge e^{r_n}. \tag{2.64}$$

The ordering of the indices r_i decides an orientations for the (sub)simplex.

The dot product of two n-forms is given by

$$\omega_r \cdot \omega_t = (\frac{1}{n!})^2 \det\big(e_{r_i} \cdot e_{t_j}\big). \tag{2.65}$$

Eq. (2.65) conforms to the standard definition of inner products for n-forms. One can check that if $e_{r_i} = e_{r_j}$ for any $i \neq j$, or if $e_{t_i} = e_{t_j}$ for any $i \neq j$, then $\omega_r \cdot \omega_t = 0$, which should hold for forms. By the definition (2.5) of the squared volume,

$$\omega_r \cdot \omega_r = (\frac{1}{n!})^2 \det\big(e_{r_i} \cdot e_{r_j}\big) = \mathbb{V}_r. \tag{2.66}$$

Vector dot products

The form representation can be used to express $a \cdot b$, $a \cdot a$, and $b \cdot b$ for the dihedral angle in terms σ_e. Let ω_h be the $d-2$-form of the hinge h, and let

$$\omega_a = \omega_h \wedge e, \quad \omega_b = \omega_h \wedge e' \tag{2.67}$$

be the $d-1$-forms of the faces of a and b (one might change the order between ω_h and e (e') if a different orientation is suitable). The edge vector e can be written as $e = a + e_\parallel$, where a is orthogonal to h and $e_\parallel$ is parallel to h. Similarly $e' = b + e'_\parallel$. Since $e_\parallel$ and $e'_\parallel$ are parallel to h, it follows from the properties of forms that $\omega_a = \omega_h \wedge a$ and $\omega_b = \omega_h \wedge b$. Therefore

$$\omega_a \cdot \omega_b = (\omega_h \wedge a) \cdot (\omega_h \wedge b) \tag{2.68}$$

$$= \frac{\omega_h \cdot \omega_h}{(d-1)^2}\, a \cdot b. \tag{2.69}$$

In the second line we used the definition (2.65) and noted that since a and b are orthogonal to h, $a \cdot e = b \cdot e = 0$ for any e of h.

Therefore

$$a \cdot b = (d-1)^2\, \frac{\omega_a \cdot \omega_b}{\omega_h \cdot \omega_h}. \tag{2.70}$$

The other terms $a \cdot a$ and $b \cdot b$ can be obtained by setting $a = b$. The numerator of (2.70) can be expressed in squared lengths using

$$\omega_a \cdot \omega_b = \frac{1}{(d-1)!^2} \det\left(e_{a_i} \cdot e_{b_j}\right) \tag{2.71}$$

$$= \frac{1}{(d-1)!^2} \det\left(\frac{1}{2}(\sigma_{0a_i} + \sigma_{0b_j} - \sigma_{a_i b_j})\right), \tag{2.72}$$

where (2.65) and (2.2) are used. Here a_i is the i-th vertex of the subsimplex a, and b_j is the j-th vertex of the subsimplex b. The vertex 0 is the one fixed when specifying the d-simplex s, and the squared lengths $\sigma_{0a_i}, \sigma_{0b_j}, \sigma_{a_i b_j}$ are inputs to simplicial gravity. According to (2.66), the denominator $\omega_h \cdot \omega_h$ of (2.70) simply equals $\mathbb{V}_h$, which is a function of squared lengths by definition (2.5) or (2.6). These formulas can then be used to express the dihedral angles in terms of squared lengths.

Incidentally, there is an alternative useful expression

$$\omega_a \cdot \omega_b = d^2 \frac{\partial \mathbb{V}}{\partial \sigma_e}, \tag{2.73}$$

where e is the edge whose vertices are outside the hinge h common to subsimplices a and b. This expression can be derived using (2.5), (2.71), (2.2) and (2.3).

2.3.5 Deficit angles

In simplicial gravity, curvature is captured by deficit angles. The deficit angle at a hinge is the difference between the flat space(time) value and the actual value for the sum of dihedral angles around the hinge.

At a hinge h in the interior of a region (instead of on the boundary), the deficit angle is defined as

$$\delta_h = 2\pi - \sum_{s \ni h} \theta_{s,h}, \tag{2.74}$$

where the sum is over all simplices s containing h.

Here 2π is the flat space(time) value. The dihedral angles around h can be obtained by projecting the simplices to the plane orthogonal to h and summing the angles around the point h projects to (Section 2.3.4). In flat Euclidean space, the angles obviously sum to 2π. In flat Lorentzian spacetime, they also sum to 2π according to Theorem 6. In the complex domain it is taken as an assumption that the flat value is 2π, so that (2.74) constitutes a definition of the complex deficit angle in general.

If the hinge h lies on the boundary of a region, the dihedral angles around it within that region can sum to less than 2π for the flat case. Suppose there are Q_h regions sharing the hinge h. Then one way to define the deficit angle is

$$\delta_h = \frac{2\pi}{Q_h} - \sum_{s \ni h} \theta_{s,h}. \tag{2.75}$$

This ensures additivity, i.e., once all the deficit angles in all regions are summed over (2.74) is recovered. Equation (2.75) is taken as the general definition of the **complex deficit angle**, with $Q_h = 1$ if h lies in the interior of the region.

2.4 Quantum gravity

Formally, gravitational path integrals take the form

$$Z = \int \mathcal{D}g \; e^{i \int d^d x \sqrt{-g}(-\lambda + kR + \cdots)} \tag{2.76}$$

in the Lorentzian, and

$$Z = \int \mathcal{D}g \; e^{\int d^d x \sqrt{g}(-\lambda + kR + \cdots)} \tag{2.77}$$

in the Euclidean. The dots stand for higher order terms that may be present. Here the Riemann tensor convention is

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma}, \tag{2.78}$$

so that as usual $\lambda > 0$ leads to a De Sitter spacetime in cosmology.

To give an exact meaning to these formal expressions non-perturbatively, one needs to specify a way to enumerate gravitational configurations to be summed over.

2.4.1 Simplicial quantum gravity

In simplicial quantum gravity,

$$Z = \int_C \mathcal{D}\sigma \; e^{E[\sigma]}, \tag{2.79}$$

where the exponent E is given below. The gravitational configurations are specified by the squared lengths σ on edges of simplicial lattices, and the path integral measure takes the form

$$\int_C \mathcal{D}\sigma = (\sum_\tau) \lim_\Gamma \prod_{e\in\Gamma} \int_{-\infty}^{\infty} d\sigma_e \; \mu[\sigma] C[\sigma]. \tag{2.80}$$

The meaning of the new symbols are explained in the next several paragraphs.

The integration measure factor $\mu[\sigma]$ is not known *a priori*. Suppose one wants to define the path integral so that even on a finite lattice (without taking the lattice refinement limit) the result is exact result. Then one idea for fixing the measure is to demand discretization independence [73]. This would lead to a non-local measure in $4D$ [74]. Alternatively, one could adopt simpler local measures and demand that the exact result be obtained only after taking the lattice refinement limit. In this case different measures could belong to a same universality class and lead to the same result in the lattice refinement limit [18]. However, there seems to be no consensus exactly which measures are correct to be used. In analogy to the continuum measures factors $(\det g)^m$, a commonly used family of simplicial measures is the product of powers of simplicial squared volumes

$$\mu[\sigma] = \prod_s \mathbb{V}_s^m \tag{2.81}$$

parametrized by m. For the Lorentzian case one could use $\mu[\sigma] = \prod_s (-\mathbb{V}_s)^m$ to make the measure positive definite, in analogy to $\prod_x (-\det g(x))^m$. When the lattice has fixed size this makes no essential difference from (2.80) since the two measures only differ by an overall constant. This can be included as a term in the integrand exponent

$$E_m = m \sum_s \log \mathbb{V}_s. \tag{2.82}$$

Any measure factor can be similarly be incorporated by setting $\mu[\sigma] = 1$ and introducing an additional term in the integrand exponent. We will adopt this formulation and fix the measure to be

$$\int_C \mathcal{D}\sigma = (\sum_\tau) \lim_\Gamma \prod_{e\in\Gamma} \int_{-\infty}^{\infty} d\sigma_e \; C[\sigma]. \tag{2.83}$$

The constraint $C[\sigma]$ specifies the integration contour and determines if the theory is for the Euclidean or Lorentzian. It equals 1 when the Euclidean/Lorentzian generalized triangle inequalities (2.12)/(2.15) are matched and vanishes otherwise. In the Lorentzian case, an additional constraint may be imposed so that each point of a simplicial manifold has two lightcones. This is explained in more detail in Section 2.4.4.

On a fixed lattice graph Γ, the gravitational configurations are summed over by integrating the squared lengths σ_e on edges e. The continuum limit $\lim_\Gamma$ is taken by going to ever finer lattice

graphs (Fig. 2.1). In practice, the lattice field theory strategy is usually adopted. Instead of taking the limit, one evaluates the path integral on a fixed graph and look for the continuum limit by searching for universality classes.

Whether topologies should be summed over in the gravitational path integral is an open question [75]. In (2.83) the sum over topologies $\sum_\tau$ is included as an option enclosed in brackets.

In (2.79) the path integral is expressed in terms of the path exponent E instead of the action S to retain unified formula for the Euclidean, Lorentzian, and general complex cases. E is related to the actions by

$$E = \begin{cases} -S^E, & \text{in Euclidean space,} \\ iS^L, & \text{in Lorentzian spacetime.} \end{cases} \tag{2.84}$$

Explicitly, E equals

$$E = \underbrace{-\lambda V}_{E_{CC}} + \underbrace{(-k)\sum_h \delta_h \sqrt{\mathbb{V}_h - 0i}}_{E_{EH}} + \underbrace{\cdots}_{E_O}. \tag{2.85}$$

E_O stands for "other terms" in addition to the cosmological constant term E_{CC} and the Einstein-Hilbert term E_{EH}. The measure factor (2.82) is an example. An R^2 term as another example is considered in Section 2.6. The terms E_{CC} and E_{EH} are discussed below.

2.4.2 Cosmological constant term

The cosmological constant term equals

$$E_{CC} = -\lambda V = -\lambda \sum_s V_s. \tag{2.86}$$

Here λ is the cosmological constant, and the sum is over all simplicial volumes $V_s = \sqrt{\mathbb{V}_s}$ as defined in (2.7).

In Euclidean space $\mathbb{V}_s > 0$, so $V_s > 0$. Therefore large volumes are suppressed by the exponent E_{CC}. This agrees with ordinary Euclidean quantum gravity. In Lorentzian spacetime $\mathbb{V}_s < 0$, so $V_s = \sqrt{\mathbb{V}_h}$ as defined in (2.7) are positive imaginary. This agrees with the usual convention for Lorentzian quantum gravity in which $E_{CC} = -i\lambda V^L$ with a positive Lorentzian volume $V^L = \sum_s |V_s|$.

2.4.3 Einstein-Hilbert term

The Einstein-Hilbert term equals

$$E_{EH} = -k \sum_h \delta_h \sqrt{\mathbb{V}_h - 0i}. \tag{2.87}$$

Here $k > 0$ is the gravitational coupling constant, the sum is over all hinges h, $\mathbb{V}_h$ is the squared volume of the hinge h, and δ_h is its deficit angle. The notation $\sqrt{z - 0i}$ is as defined in (2.32):

$$\sqrt{z - 0i} = \sqrt{r} e^{i\phi/2}, \quad z = re^{i\phi} \text{ with } \phi \in [-\pi, \pi). \tag{2.88}$$

The point is that $\sqrt{z-0i}$ is negative imaginary for $z<0$.

In the Euclidean domain $\mathbb{V}_h > 0$, so $\sqrt{\mathbb{V}_h - 0i} = \sqrt{\mathbb{V}_h} > 0$. In addition, (2.28) agrees with (2.19). Then $E_{EH} = -k\sum_h \delta_h V_h$ is minus the Einstein-Hilbert term of Euclidean simplicial quantum gravity in the convention of [18]. This in turn yields in the continuum limit

$$Z = \int \mathcal{D}g\, e^{\int d^dx\sqrt{g}(-kR)} \tag{2.89}$$

for the pure gravity path integral. Note the extra minus sign in contrast to (2.77). Since the Einstein-Hilbert term is unbounded from below, it is unclear if this sign choice is a bad one. In a follow up work, we will point out a different branch choice for the angle formula (2.28) which reproduces the the Einstein-Hilbert term with the conventional sign in the Euclidean.[6]

For a Lorentzian path integral, (2.28) is used to define the deficit angle δ_h according to (2.75). We have

$$E_{EH} = ik \sum_{h \text{ timelike}} \delta_h |V_h| - k \sum_{h \text{ spacelike}} \delta_h |V_h|, \tag{2.90}$$

where $\sum_h$ is expanded into a sum over timelike and spacelike hinges (lightlike hinges do not contribute to the exponent since $\mathbb{V}_h = 0$), and $|V_h|$ is the modulus of $V_h = \sqrt{\mathbb{V}_h}$.

Sorkin showed that

$$ik \sum_{h \text{ timelike}} \delta_h |V_h| + k \sum_{h \text{ spacelike}} \tilde{\delta_h} |V_h|, \tag{2.91}$$

reproduces $ik\int d^dx\sqrt{-g}R$ in the continuum limit when $\tilde{\delta_h}$ is positive for a spacelike Lorentz boost deficit angle [76].[7] By Theorem 4, in the convention of the present work a spacelike Lorentz boost deficit angle δ_h is negative imaginary. Therefore (2.90) also reproduces the commonly used path integral exponent $E_{EH} = iS_{EH} = ik\int d^dx\sqrt{-g}R$ of (2.76).

2.4.4 Lightcone structures

In ordinary classical space-time, each point has two lightcones attached to it. In simplicial gravity, a point can have more or fewer than two light cones (Fig. 2.7).

It is an open question whether such spacetime configurations with irregular lightcone structures should be included in the gravitational path integral. When they are included the exponent becomes complex rather than staying imaginary. This is because the constant 2π in the exponents are cancelled exactly when the angles enclose four light rays, as in ordinary flat spacetime (Theorem 7). Depending on the sign choice for the exponent, a space-time configurations with the irregular lightcone structures is either suppressed or enhanced by the additional non-vanishing real part of the exponent.

[6] I am very grateful to Bianca Dittrich and José Padua-Argüelles for discussions that clarified the sign conventions of the Einstein-Hilbert term and the mistakes I made regarding the alternatives for the Einstein-Hilbert term in a previous version of the manuscript. The discussions also clarified how one should interpret Sorkin's Lorentzian Regge action [71] so that it is holomorphic. The details of this interpretation will be reported elsewhere.

[7] In this statement R is as defined from (2.78). Note that Sorkin used an opposite sign convention for R in the original paper [76].

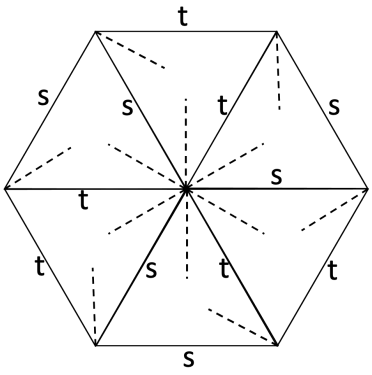

Figure 2.7: Irregular lightcone structure in $2D$. The point at the center has six light rays (dashed lines) and three lightcones, if spacelike (s) and timelike (t) edges are as assigned.

In [77], reasons are offered to prefer the enhancement (suppression) of configurations with fewer (more) than four light rays. The exponent (2.90) with the extra minus sign conforms with the opposite choice. As will be reported in details elsewhere, a different branch choice for the angle formula (2.28) reverses the enhancement/suppression. If irregular light structures are allowed in Nature, observing the enhancement/suppression effects could in principle help us to determine the branch choice.

2.5 Holomorphic flow

Analytic calculations for the non-perturbatively defined gravitational path integral is hard. In the Euclidean, one usually proceeds numerically with Markov Chain Monte Carlo simulations. The efficiency of this method relies on positivity of the path integrand in the Euclidean. In the Lorentzian, however, the path integrand is complex. The leads to the sign problem. The phase of the complex numbers summed over can fluctuate wildly to cancel each other off, which reduces the efficiency of Markov Chain Monte Carlo simulations.

The sign problem is not restricted to quantum gravity, but is also encountered in quantum theories of matter. Several methods have been developed to overcome the sign problem (see e.g., [31, 78, 79] and references therein). The basic idea of the complex path methods is to deform the integration contour to the complex to reduce the phase fluctuations. This idea is demonstrated to work for several models, including low dimension Thirring models, real time scalar field theories, and Hubbard models [31]. It has also been applied to analyze gravitational propagators for spin-foam models in the large spin limit [32].

As reviewed in [31] there are several different ways to implement the general idea of complex path deformation to overcome the sign problem. In later sections we apply the "holomorphic gradient flow" algorithm, also called the "generalized thimble" algorithm, [33, 34] to Lorentzian simplicial quantum gravity. This section summarizes the algorithm.

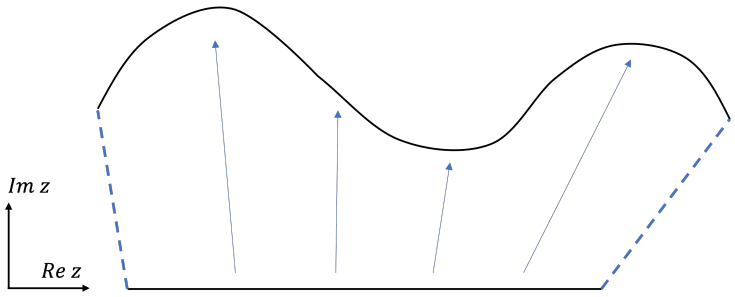

Figure 2.8: Schematic illustration of the flow region and its boundary. The original contour at the bottom is deformed into the contour at the top. The integral along these contours plus on the dashed boundaries is zero, if the function being integrated over is holomorphic inside. If the integral on the dashed boundaries are negligibly small, then the integrals on the two contours are equal up to a sign.

2.5.1 Flow equations

The celebrated Cauchy integration theorem indicates that up to a sign the integral of a complex function $f(z)$ does not change value if the integration contour is deformed through a region where $f(z)$ is holomorphic.

Cauchy's theorem admits a multi-dimensional generalization [31] which applies to path integrals of multiple variables. The holomorphic gradient flow algorithm exploits this to find deformed contours where the sign problem is mitigated. Consider a path integral with a holomorphic integrand of the form

$$Z = \int \mathcal{D}\sigma \; e^{E[\sigma]}, \tag{2.92}$$

where in $\mathcal{D}\sigma$ multiple configurations σ_e labelled by the lattice edges e are integrated over. The **flow equations** are

$$\frac{d\sigma_e}{dt} = -\overline{\partial_e E} \quad \forall e, \tag{2.93}$$

where ∂_e is a shorthand for $\frac{\partial}{\partial \sigma_e}$, and the overline stands for complex conjugation. For any point ζ in the original integration contour, the solution to (2.93) as a function of the flow time t defines the **holomorphic gradient flow** (or holomorphic flow in short) for ζ. Solving (2.93) for the whole original integration contour yields a deformation of the integration contour as a function t.

If the integral along the boundary of the flowed region is negligible, then up to a sign (2.92) can be evaluated on the flowed contour (Fig. 2.8). This could reduce the phase fluctuations for the complex numbers integrated over, because only a smaller region on the flowed contour contribute significantly to the integral, and the phase fluctuations could be small in this smaller region.

To see this, we look at the real part E_R and the imaginary part E_I of E. By (2.93),

$$\frac{\mathrm{d}E_R}{\mathrm{d}t} = \frac{1}{2}\left(\frac{\mathrm{d}E}{\mathrm{d}t} + \overline{\frac{\mathrm{d}E}{\mathrm{d}t}}\right) = \frac{1}{2}\sum_e \left(\partial_e E \frac{\mathrm{d}\sigma_e}{\mathrm{d}t} + \overline{\partial_e E \frac{\mathrm{d}\sigma_e}{\mathrm{d}t}}\right) = -\sum_e |\partial_e E|^2 \leq 0, \tag{2.94}$$

$$\frac{\mathrm{d}E_I}{\mathrm{d}t} = \frac{1}{2i}\left(\frac{\mathrm{d}E}{\mathrm{d}t} - \overline{\frac{\mathrm{d}E}{\mathrm{d}t}}\right) = \frac{1}{2i}\sum_e \left(\partial_e E \frac{\mathrm{d}\sigma_e}{\mathrm{d}t} - \overline{\partial_e E \frac{\mathrm{d}\sigma_e}{\mathrm{d}t}}\right) = 0. \tag{2.95}$$

Therefore the real part of the exponent decreases monotonically through the flow, while the imaginary part stays constant. For sufficiently long flow time, the magnitude of the integrand is exponentially suppressed for most points on the deformed contour. Only points close to the critical points of the flow obeying

$$\partial_e E = 0 \quad \forall e \tag{2.96}$$

contribute significantly.

If the phase fluctuations for such points that contribute significantly is small enough, Markov Chain Monte Carlo simulation can be efficiently performed.

2.5.2 Numerical algorithm

As a summary of Section 2.5.1, suppose:

- The holomorphic flow transverse a region where the path integrand is holomorphic;
- The boundary of the flow region have negligible contribution to the path integral.

Then the original path integral can be equally evaluated along the contour at any flow time $t = T$.

To compute the path integral on the flowed contour, one could use the holomorphic gradient flow algorithm [33, 34]. The idea is to parametrize the flowed contour by its preimage in the original contour, and perform Markov Chain Monte Carlo simulation using weights on the flowed contour. Specifically, the algorithm goes as:

1. Start with a configuration ζ in the original contour. Evolve it under the holomorphic flow by time T to obtain $\phi = \phi(\zeta)$.
2. Draw a new configuration $\zeta' = \zeta + \delta\zeta$ on the original contour, where $\delta\zeta$ is a random vector drawn from a symmetric distribution. Again evolve ζ' under the flow by time T to obtain $\phi' = \phi'(\zeta')$.
3. Accept ζ' with probability $P = \min\{1, e^{\mathrm{Re}\, E_{\text{eff}}(\phi') - \mathrm{Re}\, E_{\text{eff}}(\phi)}\}$, where E_{eff} is defined in (2.98).
4. Repeat steps 2 and 3 until a sufficient ensemble of configurations is generated.
5. Compute the expectation values using

$$\langle O \rangle = \frac{\left\langle O e^{i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, E_{\text{eff}}}}{\left\langle e^{i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, E_{\text{eff}}}}, \tag{2.97}$$

   where $\langle\cdot\rangle_{\mathrm{Re}\, E_{\text{eff}}}$ stands for the average using the ensemble just generated, and φ is defined in (2.103).

In steps 1 and 2, the evolution can be conducted through numerically integrating the ODEs (2.93). If the complexified theory has is domain on Riemann surfaces, as is the case for simplicial quantum gravity, branches need to be recorded as part of the numerical integration algorithm to make sure the system flows continuously on the Riemann surfaces. In Step 3,

$$E_{\text{eff}}(\phi) = E(\phi(\zeta)) + \log\det J(\zeta), \quad J_{ee'} = \frac{\partial \phi_e}{\partial \zeta_{e'}}, \tag{2.98}$$

where ϕ_e and ζ_e are the values ϕ and ζ take on the edge e. The Jacobian can be obtained (see Appendix A of [31]) by integrating

$$\frac{dJ_{ee'}}{dt} = \sum_{e''} \overline{H_{ee''} J_{e''e'}}, \quad H_{ee'} := -\partial_{e'}\partial_e E, \quad J_{ee'}(0) = \delta_{ee'}. \tag{2.99}$$

The function $e^{E_{\text{eff}}}$ is the integrand of the final integral to be computed, since

$$Z = \int_{M_0} e^{E(\zeta)} d\zeta \tag{2.100}$$

$$= \int_{M_T} e^{E(\phi)} d\phi \tag{2.101}$$

$$= \int_{M_0} e^{E(\phi(\zeta))} \det J \; d\zeta, \tag{2.102}$$

where we reparametrized the flowed manifold M_T by points ζ of the original manifold M_0 in the last step. Now the integrand equals $e^{E_{\text{eff}}}$ for E_{eff} defined in (2.98). Expanding E_{eff} in real and imaginary parts yields $e^{E_{\text{eff}}} = e^{\operatorname{Re} E_{\text{eff}} + i\varphi}$, where

$$\varphi = \operatorname{Im} E_{\text{eff}} = \operatorname{Im} E + \arg\det(J). \tag{2.103}$$

This explains steps 3 and 5, in which we sample (2.102) according to the magnitude $e^{\operatorname{Re} E_{\text{eff}}}$ of the integrand, and treat the phase $e^{i\varphi}$ as part of the observable in (2.97).

This algorithm can alleviate the sign problem because as $T \to \infty$, the flowed manifold approaches a combination of steepest descent contours (Lefschetz thimbles) on each of which φ is constant [31].

However, the usefulness of the algorithm is not guaranteed because of "trapping" for the Monte Carlo sampling. As noted below (2.94) $\operatorname{Re} E$ decreases monotonically under the holomorphic flow, so $\operatorname{Re} E_{\text{eff}}$ also tends to decrease. As T is increased, the probability weight $e^{\operatorname{Re} E_{\text{eff}}}$ develop peaks around the stationary points where $\partial_e E = 0$, separated by valleys where $e^{\operatorname{Re} E_{\text{eff}}}$ is exponentially suppressed. Consequently it can be hard for the Markov chain to travel across the peak regions to generate a sufficient sample.

In practice, we need to find a flow time T large enough so that the phase fluctuation in φ is sufficiently suppressed to tame the sign problem, and small enough so that the trapping of the Markov chain is sufficiently weak. More sophisticated algorithms such as the tempering algorithms [80, 81] involving multiple flow times/chains have been developed to avoid the trapping issue. In principle general Markov Chain Monte Carlo algorithms for multimodal distributions can also be applied.

2.6 2D simplicial quantum gravity

We apply the holomorphic gradient flow method to overcome the sign problem for Lorentzian simplicial gravitational path integrals. We focus on the 2D case for this initial study on the topic. The relevant expressions for the holomorphic flow equation and the Jacobian equation are given in this section. Along the way we prove a complex version of the Gauss-Bonnet theorem, which may be of independent interest. The numerical results are presented in the next section.

In $2D$, we consider the path integral

$$Z = \int \mathcal{D}\sigma \; e^{E}, \tag{2.104}$$

$$E = -\lambda V - k\sum_{v} \delta_v + a\sum_{v} \frac{\delta_v^2}{A_v} + m\sum_{t} \log \mathbb{V}_t. \tag{2.105}$$

The first (cosmological constant) and second (Einstein-Hilbert) terms are as is (2.85) specialized to $2D$. The fourth term is the measure factor term of (2.82). The third term $a\sum_v \delta_v^2/A_v$ is the R^2 term [82]. Here a is the coupling constant, and A_v is the area share of vertex v:

$$A_v = \frac{1}{3}\sum_{t\ni v} V_t = \frac{1}{3}\sum_{t\ni v} \sqrt{\mathbb{V}_t}, \tag{2.106}$$

where the sum is over triangles t containing vertex v, and $\mathbb{V}_t$ is the squared volume for triangle t calculated according to (2.5) or (2.6). The letter A instead of V is used for A_v to distinguish from the hinge (vertex in $2D$) volume $V_h = V_v$, which is usually set to 1 in $2D$.

2.6.1 Complex Gauss-Bonnet theorem

The Einstein-Hilbert term $E_{EH} = -k\sum_v \delta_v$ can actually be left out of the path integration because it is topological.

In the Euclidean domain, the celebrated Gauss-Bonnet theorem says that $E_{EH} = k2\pi\chi$, where χ is a topological invariant that is fixed by the simplicial complex, and does not depend on the particular length assignments. The same holds in the Lorentzian domain. A nice prove can be found in [71], and a slight generalization that accounts for multiple boundary components can be found in [5].

That a version of the Gauss-Bonnet theorem exists in the complex domain was suggested by Louko and Sorkin [77], but they left it as an open question to investigate.

Here we prove a complex version of the Gauss-Bonnet theorem, which generalizes the Euclidean and Lorentzian versions. It implies that on a fixed simplicial lattice, E_{EH} is constant when the Lorentzian or Euclidean contour is continuously deformed into the complex domain. Therefore E_{EH} can be taken out of the path integral in the holomorphic gradient flow algorithm.

Theorem 15 (Complex Gauss-Bonnet)**.** *On a fixed simplicial lattice, any continuous deformation of the path integration contour in the complex domain will not change the value of the Einstein-Hiblert term E_{EH}.*

If the deformation is continuously connected to the Lorentzian or the Euclidean contour,

$$E_{EH}/(-k) = 2\pi\chi, \quad \chi = V - E + T, \tag{2.107}$$

where V, E, T are the vertex, edge, and triangle numbers of the simplicial lattice, and χ is Euler number. This simple result assumes that each boundary vertex is shared by two regions.

More generally, when the numbers of regions sharing the vertices v is Q_v,

$$E_{EH}/(-k) = 2\pi\chi, \quad \chi = V^{\mathrm{o}} + \frac{1}{2}V^{\partial} - E + T + \sum_{v\in\partial}\frac{1}{Q_v}, \tag{2.108}$$

where the bulk and boundary elements are labelled by superscripts o and ∂, and the sum $\sum_{v\in\partial}$ is over all boundary vertices.

Proof. In $2D$, the Einstein-Hilbert equals

$$E_{EH}/(-k) = \sum_v \delta_v = (\sum_v 2\pi/Q_v - \sum_a \theta_a) \tag{2.109}$$

$$= (\sum_v 2\pi/Q_v - \pi N). \tag{2.110}$$

In the first line we used the definition (2.75) of the deficit angle. In δ_v for each vertex v, there is a sum over angles θ around that vertex. After $\sum_v$, we obtain a sum $\sum_a \theta_a$ is over all triangular angles of the $2D$ simplicial complex. In the second line we grouped the angles into triangles and applied Theorem 2. Here N is some integer. This shows that the E_{EH} can only take values from a discrete set labelled by N.

Under a continuous deformation of the contour, a holomorphic function such as E_{EH} can only change value continuously. Yet we just showed that the codomain of E_{EH} is a discrete set. Therefore E_{EH} cannot change value under a continuous deformation of the contour.

The claims (2.107) and (2.108) can be proved by the same argument in [71] and [5]. In the Lorentzian and Euclidean domains,

$$E_{EH}/(-k\pi) = 2V^{\mathrm{o}} + \sum_{v\in\partial}\frac{2}{n_v} - T, \tag{2.111}$$

$$0 = -2E^{\mathrm{o}} - E^{\partial} + 3T, \tag{2.112}$$

$$0 = V^{\partial} - E^{\partial}. \tag{2.113}$$

Equation (2.111) uses the fact that in the interior of the region, $Q_v = 1$, and that in the Lorentzian and Euclidean domains the angles of a triangle sum to π (Theorem 8), whence $N = T$. Equations (2.112) and (2.113) are simple facts about the simplicial lattice. Each bulk edge is shared by two faces, each boundary edge is shared by one face, and each face has three edges so (2.112) follows. The boundary is formed by a vertex-edge-vertex-edge... chain so (2.113) follows. Adding up (2.111) to (2.113) yields (2.108). Specializing to $Q_v = 2$ for all v yields (2.107). □

2.6.2 Flow equations

Because of Theorem 15, $\partial_e E_{EH} = 0$, so the flow equations (2.93) become

$$\frac{d\sigma_e}{dt} = -\overline{\partial_e E} = -\overline{\partial_e E_{CC}} - \overline{\partial_e E_{R^2}} - \overline{\partial_e E_m}. \tag{2.114}$$

For the cosmological constant term E_{CC},

$$\partial_e E_{CC} = -\lambda \partial_e V \tag{2.115}$$

$$= -\lambda \sum_t \partial_e V_t. \tag{2.116}$$

For the R^2 term E_{R^2},

$$\partial_e E_{R^2} = a \sum_v \partial_e (\frac{\delta_v^2}{A_v}) \tag{2.117}$$

$$= a \sum_v [\frac{2\delta_v \partial_e \delta_v}{A_v} - \frac{\delta_v^2 \partial_e A_v}{A_v^2}]. \tag{2.118}$$

For the measure term E_m,

$$\partial_e E_m = m \sum_t \partial_e \log \mathbb{V}_t \tag{2.119}$$

$$= m \sum_t \mathbb{V}_t^{-1} \partial_e \mathbb{V}_t. \tag{2.120}$$

Therefore

$$\frac{d\sigma_e}{dt} = -\overline{\partial_e E_{CC}} - \overline{\partial_e E_{R^2}} - \overline{\partial_e E_m} \tag{2.121}$$

$$= \lambda \sum_t \overline{\partial_e V_t} - a \sum_v \overline{(\frac{2\delta_v \partial_e \delta_v}{A_v} - \frac{\delta_v^2 \partial_e A_v}{A_v^2})} - m \sum_t \overline{\mathbb{V}_t^{-1} \partial_e \mathbb{V}_t}. \tag{2.122}$$

This formula needs to be expressed in terms of the squared lengths to be applied. While δ_v, A_v, and $\mathbb{V}_t$ in terms of the squared lengths are known from the definitions, the derivative terms in terms of the squared lengths are given below.

Volume terms

For $\partial_e V_t$ and $\partial_e A_v$, a straightforward calculation using the definitions yields

$$\partial_e V_t = \frac{\partial_e \mathbb{V}_t}{2\sqrt{\mathbb{V}_t}} = \frac{\partial_e \mathbb{V}_t}{2V_t}, \tag{2.123}$$

$$\partial_e \mathbb{V}_t = \frac{1}{8} \left(-\sigma_e + \sigma_{e1} + \sigma_{e2} \right), \tag{2.124}$$

$$\partial_e A_v = \frac{1}{3} \sum_{t \ni v} \partial_e V_t = \frac{1}{3} \sum_{t \ni v, e} \partial_e V_t, \tag{2.125}$$

where $e1, e2$ are the other two edges of the triangle t.

Angle terms

For $\partial_e \delta_v$,

$$\delta_v = 2\pi / Q_v - \sum_{t \ni v} \theta_{t,v}, \tag{2.126}$$

$$\partial_e \delta_v = -\sum_{t \ni v} \partial_e \theta_{t,v} = -\sum_{t \ni v, e} \partial_e \theta_{t,v}. \tag{2.127}$$

For a and b in triangle t meeting at vertex v, (2.26) implies

$$\frac{\partial \theta_{t,v}}{\partial \sigma_a} = \frac{\sigma_a - \sigma_b + \sigma_c}{4i\sigma_a(a \wedge b)} = \frac{\sigma_a - \sigma_b + \sigma_c}{8\sigma_a V_t}, \tag{2.128}$$

$$\frac{\partial \theta_{t,v}}{\partial \sigma_b} = \frac{-\sigma_a + \sigma_b + \sigma_c}{4i\sigma_b(a \wedge b)} = \frac{-\sigma_a + \sigma_b + \sigma_c}{8\sigma_b V_t}, \tag{2.129}$$

$$\frac{\partial \theta_{t,v}}{\partial \sigma_v} = \frac{i}{2a \wedge b} = \frac{-1}{4V_t}. \tag{2.130}$$

Here we noted that

$$a \wedge b = -2iV_t, \tag{2.131}$$

where V_t in terms of squared lengths is given in (2.11). These can be used to express (2.127) fully in the squared lengths.

2.6.3 Jacobian

The Jacobian flow equation is given in (2.99) as

$$\frac{dJ_{ee'}}{dt} = \sum_{e''} \overline{H_{ee''} J_{e''e'}}, \quad H_{ee'} := -\partial_{e'}\partial_e E, \quad J_{ee'}(0) = \delta_{ee'}. \tag{2.132}$$

Specialized to simplicial quantum gravity in $2D$,

$$H_{ee'} = -\partial_{e'}\partial_e E = -\partial_{e'}\partial_e E_{CC} - \partial_{e'}\partial_e E_{R^2} - \partial_{e'}\partial_e E_m, \tag{2.133}$$

where the Einstein-Hilbert term drop out by Theorem 15.

The cosmological constant term

The cosmological constant term is

$$\partial_{e'}\partial_e E_{CC} = -\lambda \sum_t \partial_{e'}\partial_e V_t \tag{2.134}$$

$$= -\lambda \sum_{t \ni e,e'} \partial_{e'}\partial_e V_t, \tag{2.135}$$

where it was noted that $\partial_{e'}\partial_e V_t = 0$ if the triangle t does not contain both e and e'. By (2.123) and (2.124),

$$\partial_{e'}\partial_e V_t = \frac{1}{2\sqrt{\mathbb{V}_t}}\left(\frac{-1}{2\mathbb{V}_t}\partial_e \mathbb{V}_t \partial_{e'}\mathbb{V}_t + \partial_{e'}\partial_e \mathbb{V}_t\right). \tag{2.136}$$

$$\partial_e \mathbb{V}_t = \frac{1}{8}\left(-\sigma_e + \sigma_{e1} + \sigma_{e2}\right), \tag{2.137}$$

$$\partial_{e'}\partial_e \mathbb{V}_t = \begin{cases} \frac{-1}{8}, & e = e', \\ \frac{1}{8}, & e \neq e'. \end{cases} \tag{2.138}$$

Plugging these in (2.135) yields an expression in terms of squared lengths.

Regarding computational complexity it is relevant to note that $\partial_{e'}\partial_e E_{CC}$ is quasi-local. Because the sum $\sum_{t \ni e,e'}$ in (2.135) is over triangles t that contain both e and e', if e and e' are not identical or adjacent then $\partial_{e'}\partial_e E_{CC} = 0$.

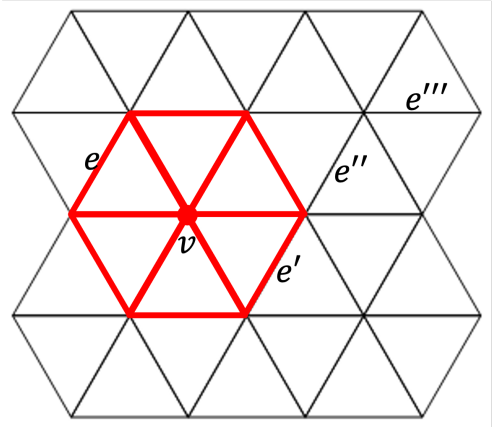

Figure 2.9: The edges that A_v and δ_v depend on are thickened. They are all within one edge away from v, and are all within two edges away from each other. A pair of edges (e.g., e and e''') more than two edges away will not find any vertex v whose A_v and δ_v depend on them both. Even a pair of edges (e.g., e and e'') two edges away may not find any vertex v whose A_v and δ_v depend on them both.

The R^2 term

For the R^2 term,

$$\begin{aligned}\partial_e\partial_{e'}E_{R^2} =& \sum_v \frac{a}{A_v^3}[2\delta_v A_v\left(-\delta_v^{(0,1)}A_v^{(1,0)} - \delta_v^{(1,0)}A_v^{(0,1)} + \delta_v^{(1,1)}A_v\right)\\ &+ \delta_v^2\left(2A_v^{(0,1)}A_v^{(1,0)} - A_vA_v^{(1,1)}\right) + 2\delta_v^{(0,1)}\delta_v^{(1,0)}A_v^2],\end{aligned} \tag{2.139}$$

where $f^{(i,j)}$ is the shorthand for $\partial_e^i\partial_{e'}^j f$.

We see that $\partial_e\partial_{e'}E_{R^2}$ is quasi-local, in the sense that $\partial_e\partial_{e'}E_{R^2} = 0$ when e and e' are more than two edges away (meaning the shortest lattice graph path touching both e and e' has more than two edges) (Fig. 2.9). This is because $\partial_e\delta_v = \partial_e A_v = 0$ if e is more than one edge away from v. If e and e' are more than two edges away, then at least one of them is more than one edge away from v for any v, whence all terms on the right hand side of (2.139) vanish.

Volume terms

By the definition of A_v,

$$A_v^{(1,0)} = \partial_e V_v = \frac{1}{3}\sum_{t\ni v}\partial_e V_t = \frac{1}{3}\sum_{t\ni v,e,e'}\partial_e V_t, \tag{2.140}$$

$$A_v^{(1,0)} = \partial_{e'} V_v = \frac{1}{3}\sum_{t\ni v}\partial_{e'} V_t = \frac{1}{3}\sum_{t\ni v,e,e'}\partial_{e'} V_t, \tag{2.141}$$

$$A_v^{(1,1)} = \partial_e\partial_{e'} V_v = \frac{1}{3}\sum_{t\ni v}\partial_e\partial_{e'} V_t = \frac{1}{3}\sum_{t\ni v,e,e'}\partial_e\partial_{e'} V_t. \tag{2.142}$$

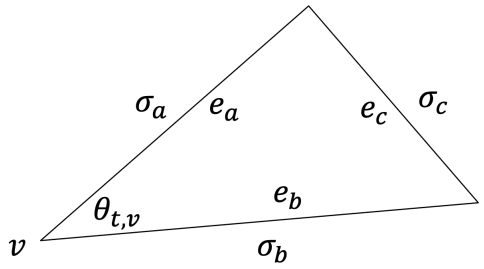

Figure 2.10: Triangle t with edges e_a, e_b, e_c whose squared lengths are $\sigma_a, \sigma_b, \sigma_c$. Edges e_a and e_b bound the angle $\theta_{t,v}$.

These can be expressed in terms of squared lengths using (2.123), (2.124), (2.136), and (2.138):

$$\partial_e V_t = \frac{\partial_e \mathbb{V}_t}{2\sqrt{\mathbb{V}_t}}, \tag{2.143}$$

$$\partial_e \mathbb{V}_t = \frac{1}{8}\left(-\sigma_e + \sigma_{e1} + \sigma_{e2}\right), \tag{2.144}$$

$$\partial_{e'}\partial_e V_t = \frac{1}{2\sqrt{\mathbb{V}_t}}\left(\frac{-1}{2\mathbb{V}_t}\partial_e \mathbb{V}_t \partial_{e'}\mathbb{V}_t + \partial_{e'}\partial_e \mathbb{V}_t\right), \tag{2.145}$$

$$\partial_{e'}\partial_e \mathbb{V}_t = \begin{cases} \frac{-1}{8}, & e = e', \\ \frac{1}{8}, & e \neq e'. \end{cases} \tag{2.146}$$

Angle terms

The terms $\delta_v^{(1,0)}$ and $\delta_v^{(0,1)}$ can be expressed in squared lengths using (2.127) - (2.130) (with labels specified in Fig. 2.10):

$$\partial_e \delta_v = -\sum_{t \ni v} \partial_e \theta_{t,v} = -\sum_{t \ni v, e} \partial_e \theta_{t,v}. \tag{2.147}$$

$$\frac{\partial \theta_{t,v}}{\partial \sigma_a} = \frac{\sigma_a - \sigma_b + \sigma_c}{4i\sigma_a(a \wedge b)} = \frac{\sigma_a - \sigma_b + \sigma_c}{8\sigma_a V_t}, \tag{2.148}$$

$$\frac{\partial \theta_{t,v}}{\partial \sigma_b} = \frac{-\sigma_a + \sigma_b + \sigma_c}{4i\sigma_b(a \wedge b)} = \frac{-\sigma_a + \sigma_b + \sigma_c}{8\sigma_b V_t}, \tag{2.149}$$

$$\frac{\partial \theta_{t,v}}{\partial \sigma_v} = \frac{i}{2a \wedge b} = \frac{-1}{4V_t}. \tag{2.150}$$

For the second derivative,

$$\delta_v^{(1,1)} = \partial_e \partial_{e'} \delta_v = -\sum_{t \ni v, e, e'} \partial_e \partial_{e'} \theta_{t,v}. \tag{2.151}$$

For e, e' ordered as e_a, e_b, e_c (Fig. 2.10), the Hessian matrix is

$$\partial_e \partial_{e'} \theta_{t,v} = \frac{1}{32V_t^3} \begin{pmatrix} \frac{X}{4\sigma_a^2} & -\sigma_c & \frac{(-\sigma_a+\sigma_b+\sigma_c)}{2} \\ -\sigma_c & \frac{Y}{4\sigma_b^2} & \frac{(\sigma_a-\sigma_b+\sigma_c)}{2} \\ \frac{(-\sigma_a+\sigma_b+\sigma_c)}{2} & \frac{(\sigma_a-\sigma_b+\sigma_c)}{2} & \frac{(\sigma_a+\sigma_b-\sigma_c)}{2} \end{pmatrix}, \tag{2.152}$$

where

$$X = \sigma_a^3 + \sigma_a^2(\sigma_c - 3\sigma_b) + 3\sigma_a(\sigma_b^2 - \sigma_c^2) - (\sigma_b - \sigma_c)^3, \tag{2.153}$$

$$Y = X(\sigma_a \leftrightarrow \sigma_b) = \sigma_b^3 + \sigma_b^2(\sigma_c - 3\sigma_a) + 3\sigma_b(\sigma_a^2 - \sigma_c^2) - (\sigma_a - \sigma_c)^3. \tag{2.154}$$

The above volume and angular terms of derivatives can be plugged into (2.139) to express it in terms of squared lengths.

The measure term

By the definition of E_m,

$$\partial_{e'}\partial_e E_m = m \sum_{t \ni e,e'} \partial_{e'}\partial_e \log \mathbb{V}_t \tag{2.155}$$

$$= m \sum_{t \ni e,e'} \frac{1}{\mathbb{V}_t^2}(\mathbb{V}_t \partial_{e'}\partial_e \mathbb{V}_t - \partial_e \mathbb{V}_t \partial_{e'} \mathbb{V}_t). \tag{2.156}$$

The previous formulas (2.124) and (2.138) can then be used to express this in terms of squared length.

2.7 Numerical results

In this section we present results of numerical simulation for the path integral

$$Z = \int \mathcal{D}\sigma \; e^E, \quad E = -\lambda V + a \sum_v \frac{\delta_v^2}{A_v} + m \sum_t \log \mathbb{V}_t, \tag{2.157}$$

parameterized by $p = (\lambda, a, m)$. These constants and the squared lengths are set unitless in this section for simplicity.

We compute the expectation value for the squared length $\langle \sigma_e \rangle = \int \mathcal{D}\sigma \; \sigma_e e^E$. According to (2.97),

$$\langle \sigma_e \rangle = \frac{\langle \sigma_e e^{i\varphi} \rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}}{\langle e^{i\varphi} \rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}}, \tag{2.158}$$

where $\langle \cdot \rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}$ is the average using the ensemble just generated, and the phase φ is the imaginary part of E_{eff}.

When φ fluctuates wildly, the sign problem is bad. The task is to choose T so that on the flowed contour the phase fluctuation is reduced. We can quantify the performance of the algorithm in alleviating the sign problem by the average phase

$$\Phi = \left| \langle e^{i\varphi} \rangle_{\mathrm{Re}\, E_{\mathrm{eff}}} \right| = \left| \frac{\int \mathcal{D}\sigma \; e^{i\varphi + \mathrm{Re}\, E_{\mathrm{eff}}}}{\int \mathcal{D}\sigma \; e^{\mathrm{Re}\, E_{\mathrm{eff}}}} \right|. \tag{2.159}$$

The closer Φ is to 1, the less the sign fluctuation, and hence the better the performance.

In the cases considered below complex contours are found where $\Phi > 0.9$.

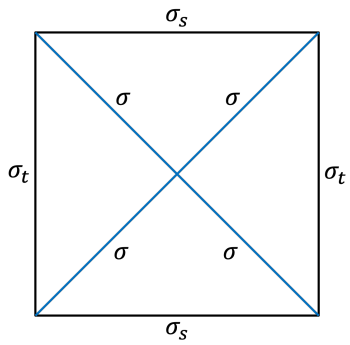

Figure 2.11: The symmetry-reduced box model with boundary squared lengths σ_t, σ_s fixed, and interior squared length σ dynamical.

2.7.1 Numerical setup

The numerical simulation is performed on a simple box model in a symmetry-reduced setting (Fig. 2.11). The boundary squared lengths are fixed at

$$\sigma_s = 1.0, \sigma_t = -1.0. \tag{2.160}$$

The four remaining edges are dynamical, and they take the same σ. In the definition of the deficit angle (2.75) we take $Q = 1$ for the interior vertex and $Q = 4$ for the boundary vertices so that the deficit angle vanishes for a box with flat geometry. At the boundary vertices A_v of (2.106) contains a sum of two triangle areas. In a different setting where the box has neighbor regions, the neighbor triangle areas would be included in the sum for A_v.

The numerical algorithm is as presented in Section 2.5.2. For any fixed flow time T, we apply the adaptive Markov Chain Monte Carlo algorithm of [83] to generate an ensemble of configurations according to the probability weight $e^{\mathrm{Re}\, E_{\text{eff}}}$. In each step we randomly pick an edge e, and propose a shift of σ_e according to a Gaussian probability distribution. The variance of the distribution is dynamical in the adaptive MCMC algorithm employed here. In this algorithm, the acceptance rate is checked every N ($N = 50$ here) steps. If the acceptance rate is below or above the target rate $r = 0.44$, the jump size is decreased or increased by

$$\delta(n) = \min(0.01, n^{-1/2}), \tag{2.161}$$

where n is the step number. That $\delta(n) \to 0$ as $n \to \infty$ ensures the asymptotic convergence of the chain.

A proposal is rejected if the Lorentzian triangle inequality is violated. In another model, one may also choose to reject a proposal if the number of light rays at a vertex is different from that of the flat configuration. However, in the symmetry-reduced box model the triangle inequality automatically implies the light ray number matching, so only the triangle inequality needs to be imposed. With this constraint, the dynamical edges can still be either timelike or spacelike.

A lower bound $E_{\text{min}} = -10.0$ is imposed on $\mathrm{Re}\, E_{\text{eff}}$ in the numerical integration for the holomorphic flow from $t = 0$ to the designated flow time $t = T$. If $\mathrm{Re}\, E_{\text{eff}}$ is too small the proposal will not be accepted. It improves the efficiency of the algorithm to simply truncate the integrator at the lower bound to move on to the next proposal.

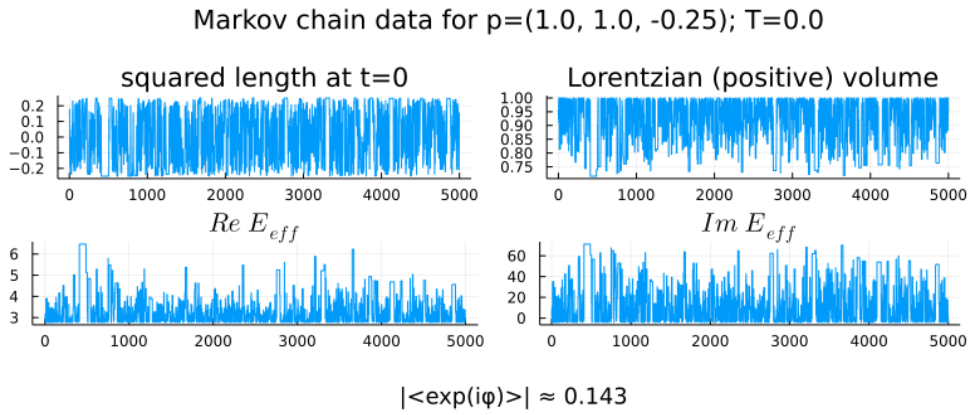

Figure 2.12: The starting case with $p = (1.0, 1.0, -0.25)$. With $T = 0.0$ the phase fluctuation is quite large.

2.7.2 Results

We consider five sets of coupling constants p. The numerical simulations are performed using the Julia programming language [84] on a personal computer. All Markov chains are obtained within about an hour. In all cases, we are able to identify a flow time T for which the sign problem is significantly ameliorated so that $\Phi > 0.9$.

Starting case

For $p = (1.0, 1.0, -0.25)$ where $m = -0.25$ for the DeWitt measure in $2D$ [18]), we consider $T = 0.0, T = 0.0005$ and $T = 0.001$ (Figure 2.12 to Figure 2.14). As the flow time T is increased from 0.0 to 0.001, the average phase Φ increases from about 0.143 to 0.929, which is close to 1 and indicates that the phase fluctuation becomes much suppressed.

Note that $\langle\sigma\rangle \approx 0$, which is not a coincidence since the model admits a $\mathbb{Z}_2$ symmetry. One can check that the transformation $\sigma \mapsto -\sigma$ on the interior squared length preserves the path integral amplitude. Therefore for any σ configuration there is the $-\sigma$ configuration with opposite contribution to $\langle\sigma\rangle$ to make $\langle\sigma\rangle = 0$ as an exact result.

On the other hand, even though the numerical estimation of $\langle\sigma^2\rangle$ is close to zero, its value is not expected to vanish. That σ^2 is small is simply because it is the square of σ which is close to zero. The third row in the figure with $T = 0.001$ shows the histograms for the real and imaginary parts of σ evaluated at the flow time T. The finite width of the distribution indicates the presence of fluctuations for the magnitude of σ.

In the following, we will change the parameters one by one to see how this influences the fluctuations reflected in the histograms and the estimated values of $\langle\sigma^2\rangle$.

Changing m

Given a new problem with a new set of parameters p, at present we do not know how to determine beforehand a suitable value of T with small enough phase fluctuation. Therefore we simply find

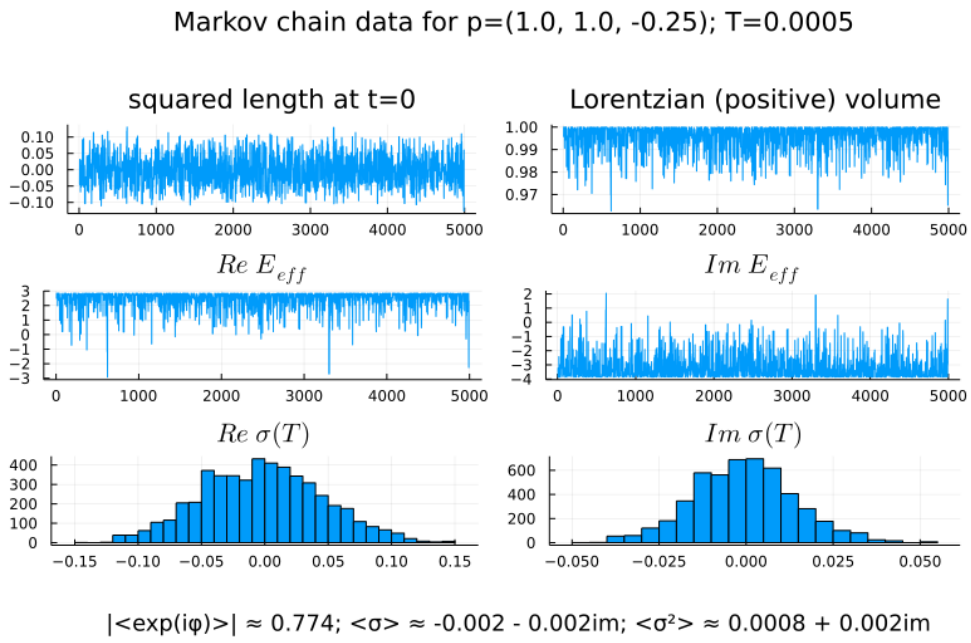

Figure 2.13: The starting case with $p = (1.0, 1.0, -0.25)$. With $T = 0.0005$ the phase fluctuation is moderately suppressed.

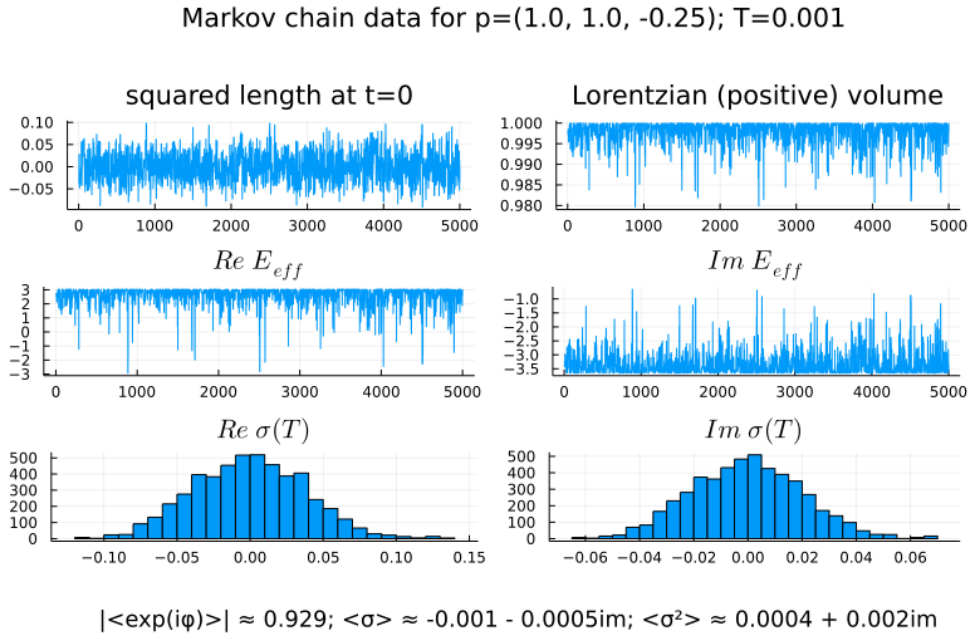

Figure 2.14: The starting case with $p = (1.0, 1.0, -0.25)$. With $T = 0.001$ the phase fluctuation is moderately suppressed.

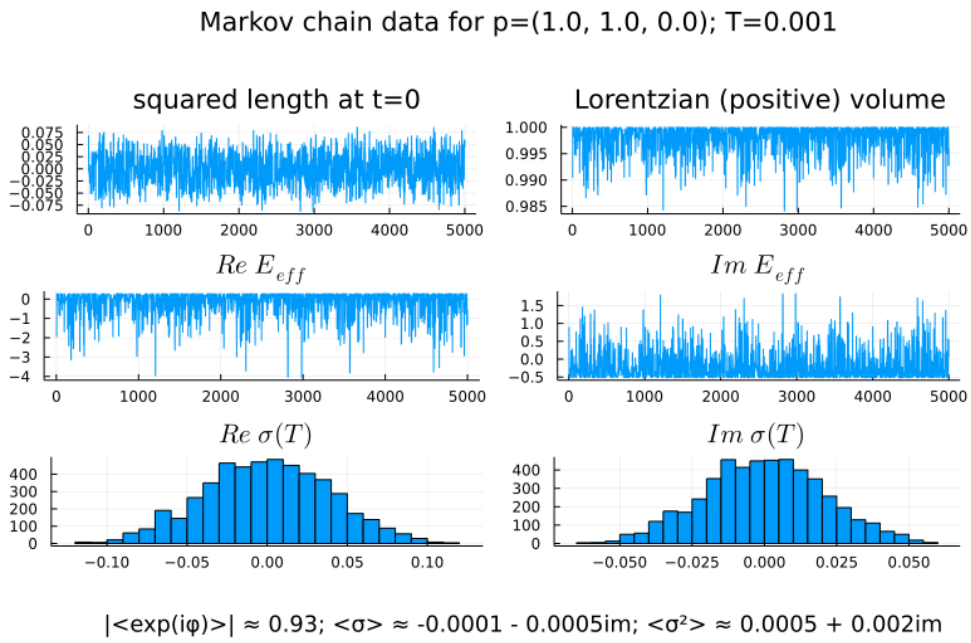

Figure 2.15: Increasing m to 0.0 does not influence the fluctuation in σ much.

a suitable value of T with $\Phi > 0.9$ by trial and error. Here and below, we directly show the results for the suitable T.

The result for m increased to 0.0 is shown in Figure 2.15. No significant difference is seen in the histogram or in $\langle\sigma^2\rangle$ in comparison to the original case of $m = -0.25$.

Changing λ

The results for λ changed to 100 and 10000 are shown in Figure 2.16 and Figure 2.17. Although it may not be so apparent from just the cases of $\lambda = 1$ and $\lambda = 100$, including the case of $\lambda = 10000$ makes it clear that the fluctuation in σ in reduced, as indicated by the decreased width of the histogram distribution and the decreased magnitude of $\langle\sigma^2\rangle$.

Changing a

The results for a changed to 100 is shown in Figure 2.18. In comparison to the cases of $a = 1$, it is quite clear that increasing a reduces the fluctuation in σ.

2.7.3 Contour boundaries

As mentioned in Section 2.5.2, to apply the holomorphic gradient flow algorithm we need that: 1) The holomorphic flow transverse a region where the path integrand is holomorphic; 2) The boundary of the flow region have negligible contribution to the path integral.

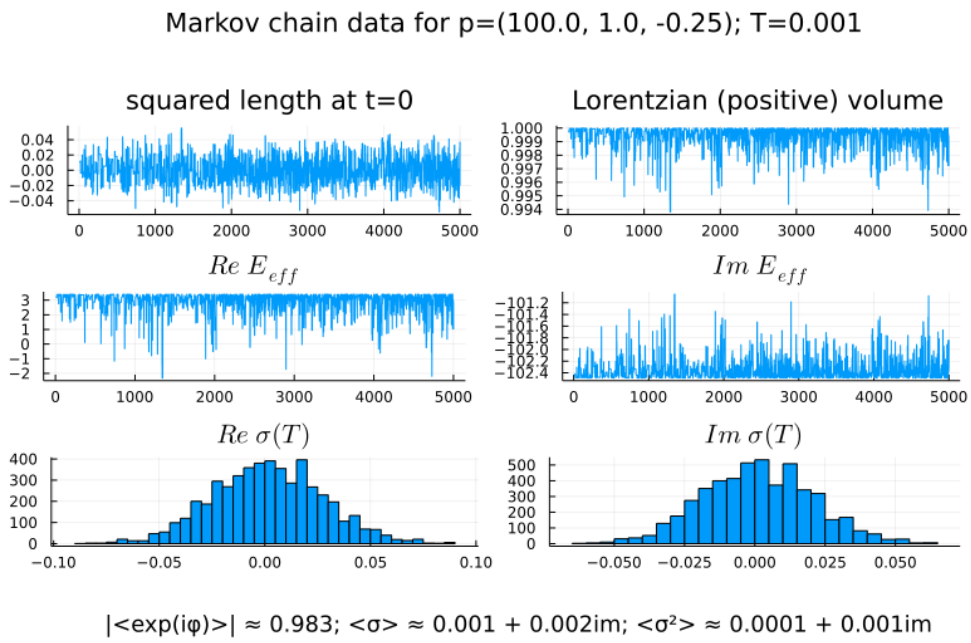

Figure 2.16: Changing λ to 100.0 slightly reduces the fluctuation in σ.

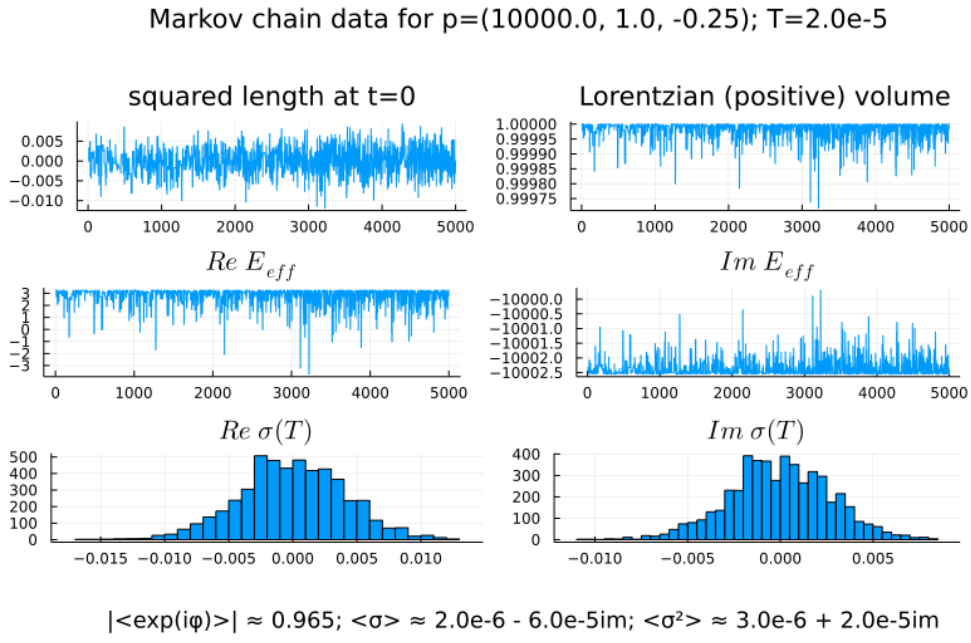

Figure 2.17: Changing λ to 10000.0 largely reduces the fluctuation in σ.

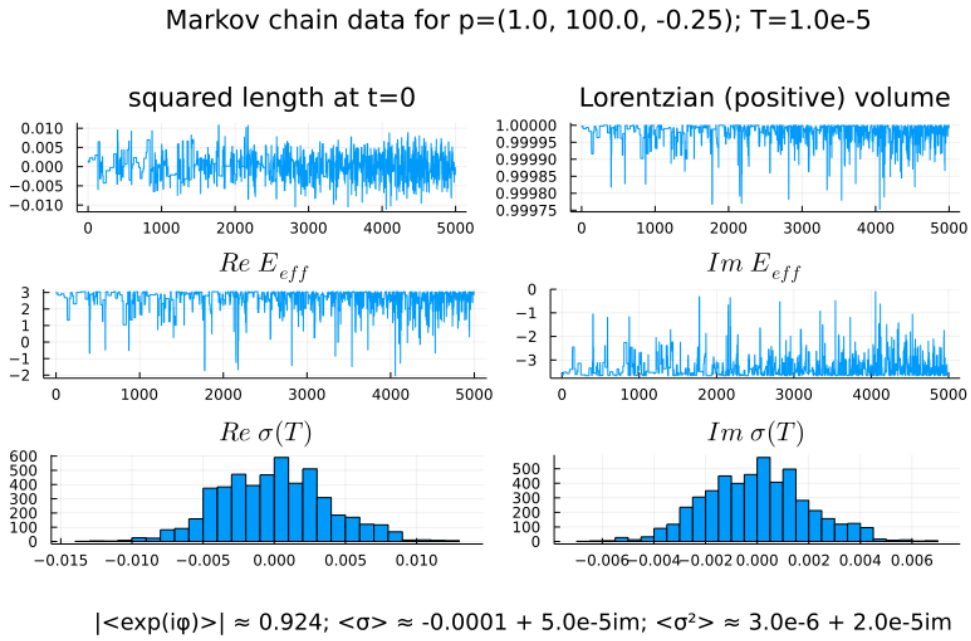

Figure 2.18: Changing a to 100.0 largely reduces the fluctuation in σ.

For simplicial quantum gravity, the boundaries are set by the branch point singularities of the path integrand, the generalized triangle inequalities, and additional constraints such as the light ray number constraint mentioned above. Within the region bounded by these boundaries, the path integrand is holomorphic, so the first requirement is met.

We now check the second requirement that the boundary of the flow region make negligible contribution to the path integral. We noted above that for the symmetry-reduced box model, the generalized triangle inequalities imply the light ray number constraint. In addition, the boundaries of the generalized triangle inequalities are set where the Lorentzian volumes vanish, i.e., $\mathbb{V}_t = 0$. Yet this coincides with one of the square root branch points singularities (see (2.11) and (2.26)). Therefore altogether we only need to consider the boundaries of the branch point singularities of the path integrand.

Along such boundaries the contribution to the path integral is infinitely suppressed. To see this, note from (2.94) that $\frac{\mathrm{d}E_R}{\mathrm{d}t} = -\sum_e |\partial_e E|^2 \leq 0$, i.e., the real part of the path exponent E decays monotonically at a rate determined by $|\partial_e E|$ along the holomorphic flow. Using the formulas of Section 2.6, one can check that $|\partial_e E| \to \infty$ at the branch point singularities. Therefore at the boundaries set by these branch points, the path integrand is infinitely exponentially suppressed. They make negligible contributions to the path integral.

2.8 Discussion

We have provided a definition of complex simplicial gravity, which reduces to Euclidean and Lorentzian simplicial gravity in special cases.

The complex formalism enabled us to perform Monte Carlo simulations for Lorentzian simplicial quantum gravity. The numerical sign problem is overcome by deforming the integration

contour into the complex.

The complex formalism also sets the path for further studies of singularity resolving processes with complex semi-classical solutions, generalizing previous studies in the symmetry-reduced setting [56, 63, 64, 65, 66, 67, 68, 69, 70, 27], and making a clear connection to the Lorentzian theory.

The numerical simulations for Lorentzian simplicial quantum gravity performed here are in a very simple setting. They are on a simple box lattice, in $1+1D$, with symmetry reduction, and for pure gravity. Future works should extend to larger lattices, higher dimensions, without symmetry reduction, and with matter coupling.

The physics theory side of these generalizations is understood. From the present work it is clear how to define complex simplicial quantum gravity on larger lattices in higher dimensions without symmetry reduction. From previous works it is clear how to couple to the matter species of the Standard Model (see e.g., Chapter 6 of Hamber's textbook [18] and references therein).

The numerics side of these generalizations still needs to be understood better. It is unclear to what extent the holomorphic gradient flow algorithm adopted here will remain efficient. Some other techniques may be needed, such as the tempered thimbles, the learnifolds, and the path optimization algorithms reviewed in [31] and further developed in, e.g., [85, 86, 87, 88].

Using the numerical tools, one could study the refinement (continuum) limit of the theory. One could investigate questions about the fate of black hole and cosmological singularities (see the Introduction section for a list of references on this topic). From a path integral perspective, if a process can be characterized by a set of path integral configurations, the formalism assigns a probability to it (which may or may not have meaning to cognitive beings such as us). Simplicial quantum gravity provides a formalism to compute and compare the probabilities for such processes.

Chapter 3

Light ray fluctuations in simplicial quantum gravity

A non-perturbative study on the quantum fluctuations of light ray propagation through a quantum region of spacetime is long overdue. Within the theory of Lorentzian simplicial quantum gravity, we compute the probabilities for a test light ray to land at different locations after travelling through a symmetry-reduced box region in 2,3 and 4 spacetime dimensions. It is found that for fixed boundary conditions, light ray fluctuations are generically large when all coupling constants are relatively small in absolute value. For fixed coupling constants, as the boundary size is decreased light ray fluctuations first increase and then decrease in a 2D theory with the cosmological constant, Einstein-Hilbert and R-squared terms. While in 3D and 4D theories with the cosmological constant and Einstein-Hilbert terms, as the boundary size is decreased light ray fluctuations just increase. Incidentally, when studying 2D quantum gravity we show that the global time-space duality with the cosmological constant and Einstein-Hilbert terms noted previously also holds when arbitrary even powers of the Ricci scalar are added. We close by discussing how light ray fluctuations can be used in obtaining the continuum limit of non-perturbative Lorentzian quantum gravity.

3.1 Introduction

The topic of light ray fluctuations is relevant to some core themes of quantum gravity.

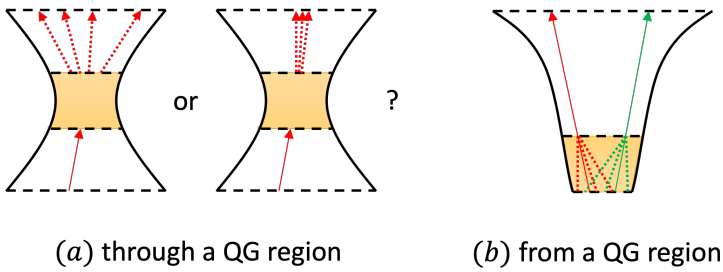

Figure 3.1: Light ray propagations affected by quantum gravitational (QG) regions.

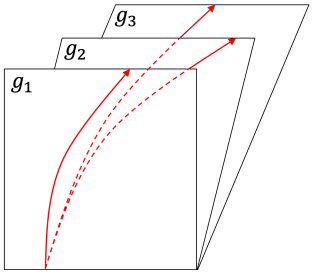

Figure 3.2: In a quantum region with superposed spacetime configurations, light rays starting at the same location on the boundary end at different locations on the other side of the boundary.

Consider the bounce scenario for quantum cosmology and quantum black holes [89] illustrated in Figure 3.1 part (a). Suppose the cosmological and/or black hole singularities in classical gravity are replaced by bouncing processes in quantum gravity. A natural question is how light rays propagate through the shaded bouncing region where quantum gravitational effects are significant. After going through the quantum gravitational region, will a light ray become quantum dispersed as in the left picture, or stay focused as in the right picture? The answer influences not just our theoretical understanding of information propagation in black hole and cosmological physics, but also experimental searches of pre-bounce relics to test the bounce scenario.

As another example consider the propagation of light rays in cosmology illustrated in Figure 3.1 part (b). Two light rays that were never in causal contact if spacetime was treated classically (they trace out the solid lines in the figure) could actually have been in causal contact if spacetime is treated quantumly to allow quantum fluctuations of the light ray paths (dashed lines). This difference could affect our qualitative and quantitative understanding on early universe cosmology [90], in particular on the horizon problem and the inflation hypothesis.

In general, we are interested in the propagation of light rays across of a region of quantum spacetime where different spacetime configurations are in superposition, yielding different paths of light ray propagation (Figure 3.2). Since strong gravity is involved in the cosmology and black hole scenarios mentioned above, we are interested in a non-perturbative treatment of the problem.

Conceptually, it is very clear how to study the problem in non-perturbative Lorentzian gravitational path integrals. Let there be a region of quantum spacetime with fixed boundary configurations. Different spacetime configurations compatible with the boundary configuration are summed over, yielding amplitudes for the light ray to land at different locations.

Practically, how smoothly the study would proceed depends very much on which non-perturbative Lorentzian gravitational path integral is used. In spin-foam models, before proceeding it needs to be clarified whether the spin-foams represent continuum spacetime configurations or some fundamentally discrete structure. This choice affects where light rays and causal paths [91] can travel on a spin-foam (see [92] for a related discussion). In quantum causal set path integrals, it needs to be decided if the path integral should be restricted to configurations corresponding to a particular spacetime dimension, and if so how (see Section 6.4 of [35] for a discussion). In causal dynamical triangulation, the topic is more accessible since the light ray path on a piecewise flat spacetime configuration is obtainable and the path integral is well specified in different dimensions [28, 29]. While we think it is possible to study the topic under discussion in causal dynamical triangulation, to our knowledge such studies have not been carried out before.

In this work we study light ray fluctuations across a quantum region of spacetime in Lorentzian simplicial quantum gravity [2]. The study of gravitational path integrals defined in terms of simplicial spacetime configurations á la Regge [14] has a long history [15, 16, 17, 18, 19]. While previous works focused on the Euclidean theory, there has been a growing interest in the Lorentzian theory in recent years [23, 24, 26, 27, 5, 2, 93]. As in causal dynamical triangulation, the light ray paths on a piecewise flat simplicial spacetime configuration is obtainable, in particular using Lorentzian trigonometry which we illustrate in Section 3.3.1. From there one could integrate over spacetime configurations corresponding to different light ray landing locations to obtain the quantum amplitudes and the probabilities.

The non-perturbative Lorentzian path integral is not easy to compute due to the complex phase of the integrand. To facilitate the study we make two simplifications. First, we consider the fluctuation of *test* light rays. In other words we consider models of pure quantum gravity to infer the light paths from the gravitational configurations alone, without introducing matter degrees of freedom. This means the backreaction of light on gravity is not taken into account in this simplified study. Second, we focus on a symmetry-reduced "box model" with simple boundary conditions and only one dynamical degree of freedom. This allows us to evaluate the Lorentzian path integral through direct numerical integration.

In this simplified model, we ask how light ray fluctuations are affected by the coupling constants and the size of the region determined by the boundary conditions in 2,3 and 4 spacetime dimensions. For fixed boundary conditions, we find that light ray fluctuations are generically large when all coupling constants are relatively small in absolute value in all dimensions. For fixed coupling constants, we find that as the boundary size is decreased, light ray fluctuations first increase and then decrease in a 2D theory with the cosmological constant, Einstein-Hilbert and R^2 terms. On the other hand, as the boundary size is decreased light ray fluctuations just increase in 3D and 4D theories with the cosmological constant and Einstein-Hilbert terms. As a side result, when studying 2D quantum gravity we show that the global time-space duality with the cosmological constant and Einstein-Hilbert terms noted previously [5] also applies when arbitrary even powers of the Ricci scalar are added.

These results point to light ray fluctuations as potentially useful in the study of the renormalization group and the continuum limit of non-perturbative Lorentzian quantum gravity. In performing renormalization group type analysis by refining the lattices to approach the continuum limit, it is important to find physical quantities to compare across different lattices. The present study reveals the light ray amplitudes and light ray probabilities as candidate physical quantities to compare across lattices. These quantities which are not accessible in the Euclidean offer some new opportunities to be explored in future works.

The paper is organized as follows. The formalism of Lorentzian simplicial quantum gravity is introduced in Section 3.2. The symmetry-reduced box model and the formulas for light ray locations across the box region are presented in Section 3.3. The results in 2,3 and 4D are presented in Section 3.4 to Section 3.6. The important quantity of the light ray amplitude and its relevance to the continuum limit of Lorentzian quantum gravity is discussed in Section 3.7. A brief summary including discussions on future prospects is given in Section 3.8.

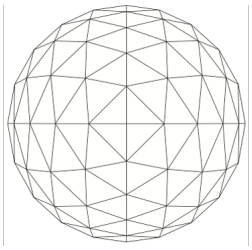

Figure 3.3: Describing curved space/spacetime by gluing flat simplicies.

3.2 Lorentzian simplicial quantum gravity

Formally, gravitational path integrals take the form

$$Z = \int \mathcal{D}g \; A[g] \tag{3.1}$$

of a sum over gravitational configurations g weighted by amplitudes $A[g]$. To fully define the path integral, we need specify a way to enumerate the gravitational configurations to perform the sum. In simplicial quantum gravity [14, 15, 16, 17, 18, 19], the sum is over simplicial spacetime configuration which describe curved spacetime by combining flat simplicies (Figure 3.3). While extensive works have been carried out in the past in the Euclidean signature, the present study of light ray fluctuations is based on a Lorentzian version of the theory [2] that sums Lorentzian simplicial spacetime configurations [94, 71].

In classical General Relativity a spacetime configuration is characterized by the metric field g_{ab}. Its physical meaning is that the line element $ds^2 = g_{ab}dx^a dx^b$ indicates the squared length between infinitesimally separated points. In simplicial gravity a spacetime configuration is characterized by the **squared lengths**

$$\sigma_e = \int_e ds^2 \tag{3.2}$$

integrated along the simplicial edges e. As such, σ_e is the finite version of the line element ds^2. In the metric signature

$$(-,+,\cdots,+) \tag{3.3}$$

used here, σ_e can be smaller than, equal to, or greater than zero, corresponding to the edge being timelike, lightlike, and spacelike. The spacetime within a simplex is taken to be flat, and the simplicial configuration is fully characterized by σ on all the edges.

The gravitational path integral is then a sum over simplicial spacetime configurations specified by the edge squared lengths on simplicial lattice graphs Γ:

$$Z = \int \mathcal{D}\sigma \; e^{E[\sigma]}, \tag{3.4}$$

$$\int \mathcal{D}\sigma = \lim_{\Gamma} \prod_{e\in\Gamma} \int_{-\infty}^{\infty} d\sigma_e \; \mu[\sigma]L[\sigma]C[\sigma]. \tag{3.5}$$

On a finite lattice Γ the integral gives an approximate result. The exact result is approached in the limit $\lim_\Gamma$ of infinitely refining the lattice.[1] The symbols $L[\sigma], C[\sigma], \mu[\sigma], E[\sigma]$ stand for the Lorentzian constraint, the lightcone constraint, the integration measure factor, and the path integral exponent. Their forms are specified below.

3.2.1 Path integral exponent

The formal continuum path integral exponent takes the form

$$E = i \int d^D x \sqrt{-g}(-\lambda + kR + aR^2 + \cdots), \tag{3.6}$$

where λ, k, a are coupling constants and $\cdots$ signifies the possibility of including additional terms. We want to find the simplicial versions of the exponent.

2D

In 2 spacetime dimensions the simplicial path integral exponent takes the form

$$E = -\lambda \sum_t A_t - k \sum_v \delta_v + a \sum_v \frac{\delta_v^2}{A_v} + \cdots . \tag{3.7}$$

The first term is the cosmological constant term. For a triangle t with squared edge lengths $\sigma_{01}, \sigma_{02}, \sigma_{12}$, the squared area formula

$$\mathbb{A}_t = \frac{1}{16}\left(-\sigma_{01}^2 - \sigma_{02}^2 - \sigma_{12}^2 + 2\sigma_{01}\sigma_{02} + 2\sigma_{01}\sigma_{12} + 2\sigma_{02}\sigma_{12}\right) \tag{3.8}$$

generalizes Heron's formula to apply to both Euclidean and Lorentzian cases. $\mathbb{A}_t$ is positive in the Euclidean and negative in the Lorentzian, just like the squared area factor $g = \det g_{ab}$. The Lorentzian triangle area

$$A_t = \sqrt{\mathbb{A}_t} \tag{3.9}$$

is positive imaginary and forms the analogue of $\sqrt{g} = i\sqrt{-g}$. This explains why $-\lambda \sum_t A_t$ is the simplicial version of $-i\lambda \int d^D x \sqrt{-g}$.

The second term is the Einstein-Hilbert term. Simplicial gravity is based on the idea that composing flat simplicies can describe curved spacetime configurations. This is possible because the sum of angles around a vertex can differ from the flat spacetime value. The difference is encoded in the **deficit angle**

$$\delta_v = F_v - \sum_{t \ni v} \theta_{t,v}. \tag{3.10}$$

In this formula, $\sum_{t \ni v} \theta_{t,v}$ is the sum of triangle angles $\theta_{t,v}$ around a vertex v, while F_v is flat spacetime value. If v lies in the interior of a region, then as explained below $F_v = 2\pi$. If v lies

[1] It is possible to include an additional sum over spacetime topologies by summing over lattices with different topologies in (3.5). Yet in the simple box models studied below we will focus on spacetime configurations with the trivial topology.

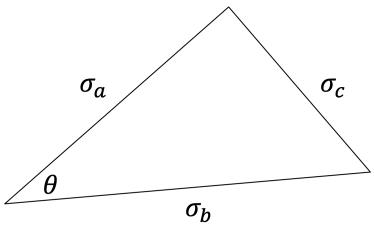

Figure 3.4: Angle of a Lorentzian triangle.

on the boundary of a region, then F_v depends on how many pieces of spacetime regions share the vertex v, and can only be fixed on a case by case basis. To obtain a lattice version of the Ricci scalar R we even out δ_v to unit areas through $-\delta_v/A_v$ where

$$A_v = \frac{1}{3}\sum_{t\ni v} A_t. \tag{3.11}$$

is the vertex share of the Lorentzian area. A triangle t contains three vertices v, so a vertex shares $1/3$ of the triangle's area A_t. In the continuum limit $-\delta_v/A_v$ approaches R for θ defined in (3.12) [94, 2].[2] This explains why $-k\sum_v A_v\frac{\delta_v}{A_v} = -k\delta_v$ forms the lattice version of the Einstein-Hilbert term $ik\int d^Dx\sqrt{-g}R$. In 2D a Lorentzian version of the Gauss-Bonnet theorem [71] indicates that the Einstein-Hilbert term is a topological invariant. Therefore it can usually be taken out of the path integration to simplify the study.

The third term $a\sum_v \frac{\delta_v^2}{A_v} = aA_v(-\sum_v \frac{\delta_v}{A_v})^2$ is the lattice version of $a\int d^Dx\, i\sqrt{-g}R^2$, given that $-\delta_v/A_v$ is the lattice analogue of R. This term is not a topological invariant and allows non-trivial stationary points for the action. In this work, we focus on these first three terms in the exponent for the studies.

Since the integration variable is the squared length, we need to express the deficit angle δ_v and hence the triangle angles $\theta_{t,v}$ in squared length. For a Lorentzian or Euclidean triangle with squared lengths as shown in Figure 3.4 [2],

$$\theta = -i\,\mathrm{Log}\,\alpha, \tag{3.12}$$

$$\alpha = \frac{\sigma_a+\sigma_b-\sigma_c+\sqrt{\sigma_a^2+\sigma_b^2+\sigma_c^2-2\sigma_a\sigma_b-2\sigma_b\sigma_c-2\sigma_c\sigma_a}}{-2\sqrt{-\sigma_a}\sqrt{-\sigma_b}}, \tag{3.13}$$

where Log stands for the principle branch of the log function ($\mathrm{Im}\,\mathrm{Log}\,\alpha = \pi$ for $\alpha < 0$). To understand this formula, consider for a moment an Euclidean triangle with squared lengths $\sigma_a, \sigma_b, \sigma_c$. In this case the first part of α, $\frac{\sigma_a+\sigma_b-\sigma_c}{-2\sqrt{-\sigma_a}\sqrt{-\sigma_b}}$, is simply $\cos\theta$ according to the law of cosines. The second part, $\frac{\sqrt{\sigma_a^2+\sigma_b^2+\sigma_c^2-2\sigma_a\sigma_b-2\sigma_b\sigma_c-2\sigma_c\sigma_a}}{-2\sqrt{-\sigma_a}\sqrt{-\sigma_b}}$, is simply $i\sin\theta$ by recognising the numerator as $4i$ times the triangle area according to Heron's formula. Therefore $\theta = -i\,\mathrm{Log}\,\alpha$ holds for an Euclidean triangle. By allowing the squared lengths to be negative we arrive at the angle formula (3.12) which applies in both the Lorentzian and the Euclidean.

It can be check that in flat spacetime the angles around a vertex sum to 2π [2], which confirms the claim above about F_v.

[2] Here the sign convention for the curvature is set by $R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}$.

Higher dimensions

In higher dimensions, we consider the simplicial path integral exponent

$$E = -\lambda \sum_s \sqrt{\mathbb{V}_s} + ik \sum_h \delta_h \sqrt{-\mathbb{V}_h} + \cdots . \tag{3.14}$$

The first term is the cosmological constant term. For a D-simplex s with squared edge lengths $\sigma_{01}, \sigma_{02}, \cdots$, the Cayley-Menger determinant

$$\mathbb{V}_s = \frac{(-1)^{D+1}}{2^D (D!)^2} \begin{vmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & \sigma_{01} & \sigma_{02} & \dots & \sigma_{0d} \\ 1 & \sigma_{01} & 0 & \sigma_{12} & \dots & \sigma_{1d} \\ 1 & \sigma_{02} & \sigma_{12} & 0 & \dots & \sigma_{2d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_{0d} & \sigma_{1d} & \sigma_{2d} & \dots & 0 \end{vmatrix} . \tag{3.15}$$

yields its squared volume, which generalizes (3.8) to D dimensions. $\mathbb{V}_s$ is the analogue of g and both are positive in the Euclidean and negative in the Lorentzian. This explains why $-\lambda \sum_s \sqrt{\mathbb{V}_s}$ is the simplicial version of $-i\lambda \int d^D x \sqrt{-g}$.

The second term is the Einstein-Hilbert term. Imagine we generate a D-dimensional simplicial configurations by extending a 2D Lorentzian configuration uniformly in D-2 additional spatial dimensions. The Einstein-Hilbert term would be $-k \sum_v \delta_v V$, which is the 2D term $-k \sum_v \delta_v$ times $V > 0$, the spatial volume extended in the D-2 additional dimensions. Given a general D-dimensional Lorentzian simplicial configuration, we could imagine that this is arrived at by non-uniformly extending a 2D configuration. The Einstein-Hilbert term would be

$$ik \sum_h \delta_h \sqrt{-\mathbb{V}_h}. \tag{3.16}$$

Here the sum is over codimension 2 subsimplices (edges in 3D, triangles in 4D etc.) referred to as **hinges** and labelled by h. At each hinge the deficit angle δ_h is obtained by projecting the D-simplices containing the hinge to the 2D plane orthogonal to the hinge and computing the deficit angle at the vertex where the hinge projects to in this plane (see [2] for the formula of the deficit angle in terms of squared lengths). The plane can be Euclidean or Lorentzian, and (3.10), (3.12), (3.13) apply equally well given the squared lengths. As suggested above by the multiplication by V, the hinges are extended non-uniformly in D-2 dimensions so we multiply hingewise by the volumes $-i\sqrt{-\mathbb{V}_h}$, where $\mathbb{V}_h$ is the squared volume defined by (3.15) which applies to both Euclidean and Lorentzian hinges. This explains $ik \sum_h \delta_h \sqrt{-\mathbb{V}_h}$ as the Einstein-Hilbert term for the extended configuration, which has the correct continuum limit [94, 2].

In contrast to 2D, in higher dimensions the Einstein-Hilbert term is not a topological invariant, so we will ignore higher order terms in the following given that the first two terms already yields a non-trivial theory.

3.2.2 Measure factor

For the integration measure factor $\mu[\sigma]$, a commonly used family of local measures is the product of powers of the simplicial (square) volumes

$$\mu[\sigma] = \prod_s \mathbb{V}_s^m \tag{3.17}$$

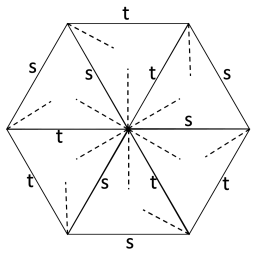

Figure 3.5: With s representing spacelike edges and t representing timelike edges, the vertex at the center of the figure has six light rays (dashed lines) and three lightcones.

parametrized by $m \in \mathbb{R}$ [18].[3] In this work we adopt (3.17) as the measure factor. The constant m, like the coupling constants λ, k and a, is treated as a parameter of the theory.

3.2.3 Lorentzian and lightcone constraints

Without the Lorentzian constraint $L[\sigma]$, the integral (3.5) also includes non-Lorentzian configurations such as Euclidean ones where all edges have positive square lengths and all simplices have positive squared volumes. The Lorentzian constraint $L[\sigma]$ serves to ensure that only Lorentzian spacetime configurations are included in (3.5).

A simplex s is Lorentzian, i.e., embedable in 2D Minkowski spacetime, if and only if [24, 26]

$$\mathbb{V}_s < 0; \text{ and } \mathbb{V}_r < 0 \implies \mathbb{V}_t \leq 0 \text{ for all } t \supset r. \tag{3.18}$$

This says that: 1) The simplex s itself has negative squared volume, and 2) if a subsimplex r of s has negative squared volume, then all higher-dimensional subsimplices t that contain r have non-negative squared volumes.

The first condition is easy to digest since we know that the metric determinant g as the infinitesimal squared volume is negative in the Lorentzian and positive in the Euclidean. The second condition is there because a timelike subsimplex cannot be embedded in a higher dimensional spacelike subsimplex (e.g., for a 4-simplex s, a timelike triangle subsimplex r cannot be embedded in a spacelike tetrahedron subsimplex t), and the first condition is not enough to ensure this. The Lorentzian constraint is then

$$L[\sigma] = \begin{cases} 1, & \text{if (3.18) holds for all simplices } s, \\ 0, & \text{otherwise.} \end{cases} \tag{3.19}$$

The Lorentzian constraint is not enough to ensure the path integral includes only ordinary spacetime configurations. In an ordinary spacetime configuration, each point has two lightcones attached to it. Without the lightcone constraint $C[\sigma]$, the integral (3.5) can include configurations where a vertex can have fewer or more than two lightcones, with the latter case illustrated in the

[3] Since $\mathbb{V}_s$ is negative in the Lorentzian, one could use $\mu[\sigma] = \prod_s (-\mathbb{V}_s)^m$ instead to make the measure positive. However, this makes no essential difference on a fixed lattice because the two measures differ at most by an overall constant.

2D configuration of Fig. 3.5. The lightcone constraint

$$C[\sigma] = \begin{cases} 1, & \text{all interior points have two lightcones attached,} \\ 0, & \text{otherwise} \end{cases} \tag{3.20}$$

ensures that all points in the interior region of the path integral configurations have two lightcones attached.

Since each simplex is just a portion of Minkowski spacetime, the lightcone number is always 2 in the interior of the simplices. On a boundary point of a simplex where multiple simplices meet, we need to count lightcones for each simplex and add up the number to check the lightcone constraint. For instance the center vertex in the 2D configuration of Fig. 3.5 is met by six triangles each coming with a light ray. Hence there are 3 lightcones, violating the lightcone constraint.

3.2.4 Scaling identity

Later we will study the dependence of the path integral on the boundary edge squared lengths. For this purpose it is useful to derive a scaling identity that relates a scaling of the boundary condition to a scaling of the coupling constants. Consider the path integral on a fixed lattice graph[4] Γ with fixed boundary squared lengths σ_B

$$Z[\sigma_B, c_i, m] = \prod_{e\in\Gamma} \int_{\sigma_B} d\sigma_e \; L[\sigma]C[\sigma] \; (\prod_s \mathbb{V}_s^m) e^{\sum_i c_i E_i}, \tag{3.21}$$

where the exponent is expressed with the coupling constants c_i of length dimension d_i. Rewriting σ on all edges as $l^2\sigma'$ where l is an arbitrary constant number yields

$$\begin{aligned} Z[\sigma_B, c_i, m] =& \prod_{e\in\Gamma} \int_{l^2\sigma'_B} d(l^2\sigma'_e) \; L[l^2\sigma']C[l^2\sigma'] \; (\prod_s (l^{2D}\mathbb{V}'_s)^m) e^{\sum_i c_i l^{-d_i} E'_i} && (3.22) \\ =& l^{2N_e+2mDN_s} \prod_{e\in\Gamma} \int_{\sigma'_B} d\sigma'_e \; L'C' \; (\prod_s (\mathbb{V}'_s)^m) e^{\sum_i c_i l^{-d_i} E'_i} && (3.23) \\ =& l^{2N_e+2mDN_s} \; Z[l^{-2}\sigma_B, c_i l^{-d_i}, m]. && (3.24) \end{aligned}$$

Here D is the spacetime dimension, N_e is the number of non-boundary edges in Γ, N_s is the number of simplices, and f' is the shorthand for $f[\sigma']$ for a function $f[\sigma]$. In the second line we noted that the constraints L and C take the same values for σ and σ'.

Setting $l^2\sigma_B$ in place of σ_B for (3.24), we get that

$$Z[l^2\sigma_B, c_i, m] = l^{2N_e+2mDN_s} \; Z[\sigma_B, c_i l^{-d_i}, m]. \tag{3.25}$$

This says that changing the boundary condition by an arbitrary factor l^2 (LHS) can equivalently be achieved by scaling the coupling constants while keeping the boundary condition fixed (RHS).

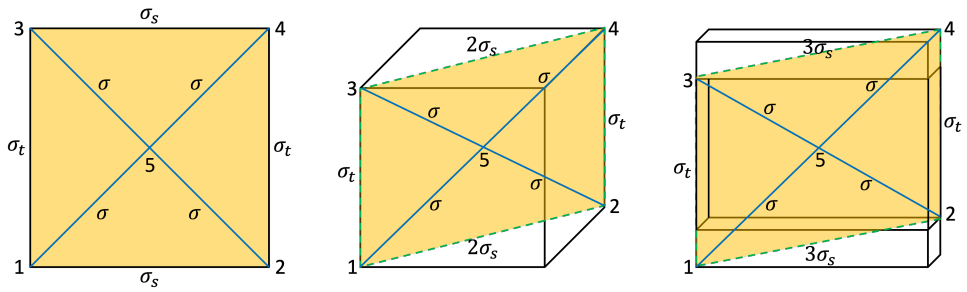

Figure 3.6: Symmetry-reduced box models in 2,3 and 4 spacetime dimensions. The diagonal 2D plane where the light ray travels is shaded and its boundary and interior squared lengths are labelled.

3.3 Symmetry-reduced box model

A box model is a hypercube with an interior vertex connected to each of the boundary vertices by an edge. The hypercube is thus divided into flat hyperpyramids formed by the interior vertex as the tip and a face of the hypercube as the base. Here we focus on symmetry-reduced box models in 2,3 and 4 spacetime dimensions illustrated in Figure 3.6. The boundary condition is that:

$$\begin{aligned} &\text{All boundary spacelike edges have squared length } \sigma_s > 0; \\ &\text{All boundary timelike edges have squared length } \sigma_t < 0. \end{aligned} \tag{3.26}$$

In a path integral configuration all the interior squared lengths take the same value σ and in this sense the model is symmetry-reduced.

In computing the path integral exponent, we assume the box to be a standalone region without neighbors. Therefore in 2D the vertex area A_v of (3.11) is the sum of two triangles. In addition, F_v in the definition of the deficit angle (3.10) is fixed to take the value for a flat hypercube in general dimensions. For example in 2D $F_v = \pi/2$ as a quarter of the full angle 2π.

Given that the hyperpyramids are not simplices, they need to be divided into simplices before applying the formulas of Section 3.2.1. Alternatively the flat hyperpyramids could be treated as the elementary building blocks that describe curved spacetime configurations by having non-vanishing deficit angles where they are glued together. Since the hyperpyramids are flat, in the simplex description their interior deficit angles vanish. Hence the simplex description agrees with the pyramid description on the Einstein-Hilbert term. Since non-squared Lorentzian and Euclidean volumes are both additive, the two descriptions also agree on the cosmological constant term. The only difference lies in the measure term, which in the simplex description is given by Equation (3.17) as $\mu[\sigma] = \prod_s \mathbb{V}_s^m$. In the pyramid description this is replaced by

$$\mu[\sigma] = \prod_p \mathbb{V}_p^m, \tag{3.27}$$

where the sum is over hyperpyramids p, and the hyperpyramid squared volume $\mathbb{V}_p$ is negative for Lorentzian hyperpyramids. $\mathbb{V}_p$ equals $-(\sum_{s\in p}\sqrt{-\mathbb{V}_s})^2$ in terms of the simplex decomposition of the hyperpyramid. Below we conform with the pyramid description for the integration

[4]If one considers lattice refinement to take the continuum limit, then the scaling identity (3.24) may receive modifications for anomalous scaling dimensions. Although a brief schematic discussion on the continuum limit is given in Section 3.7, the computations performed in this work are based on fixed lattices so we will not consider anomalous dimensions in the scaling identity.

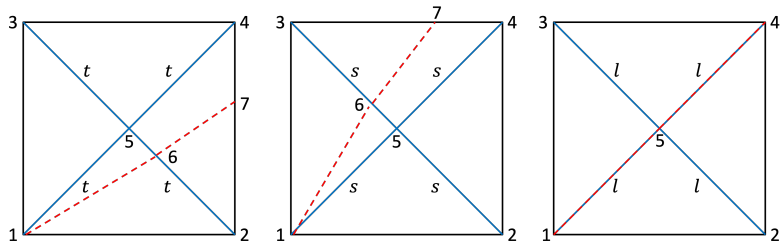

Figure 3.7: Situations for the light ray (dashed lines) emanating from vertex 1 in the symmetry-reduced box model when the interior edge is timelike (t), spacelike (s), and lightlike (l).

measure term, because the integrated variables are the squared edges on the boundaries of the hyperpyamid instead of the simplices.

3.3.1 Light ray locations

In all dimensions, we consider a light ray travelling through the quantum spacetime region within the diagonal 2D planes shaded in Figure 3.6, again starting from vertex 1. In 2D, the spacelike edges of the plane have squared length σ_s. In 3D, the spacelike edges of the plane have squared length $2\sigma_s$ because they are the diagonal edges of the 2D base squares. In 4D, the spacelike edges of the plane have squared length $3\sigma_s$ because they are the diagonal edges of the 3D base cubes.

Therefore in all dimensions, the light ray travels through a 2D simplicial configuration consisting 4 triangles as illustrated in Figure 3.7. Depending on whether the interior edge is timelike or spacelike, the light ray will land on either edge 24 or edge 34. If the interior edge is lightlike, the light ray will land right on the vertex 4.

To find the precise light ray landing location we apply Lorentzian trigonometry to triangles with one lightlike edge. Consider a triangle with squared lengths σ_a, σ_b and σ_c where edge c is lightlike. For the angle bounded by the edges a and b, setting $\sigma_c = 0$ in (3.13) yields

$$\alpha^2 = \left(\frac{\sigma_a + \sigma_b + \sqrt{\sigma_a^2 + \sigma_b^2 - 2\sigma_a\sigma_b}}{-2\sqrt{-\sigma_a}\sqrt{-\sigma_b}}\right)^2 = \left(\frac{\sigma_a + \sigma_b + |\sigma_a - \sigma_b|}{-2\sqrt{-\sigma_a}\sqrt{-\sigma_b}}\right)^2 = \frac{\sigma_+}{\sigma_-}, \tag{3.28}$$

where σ_+ is the larger real number between σ_a and σ_b, while σ_- is the smaller one ($\sigma_a = \sigma_b$ is impossible because $\mathbb{A}_t$ would be zero which violates the Lorentzian constraint of Section 3.2.3).

For a box with timelike interior edges illustrated in the left of Figure 3.7, $\sigma_{27} < \sigma_{26} < 0 < \sigma_{12}$. The first inequality holds because the future-pointing light ray 67 is moving "upwards" such that edge 27 has a larger timelike length than edge 26, which implies $\sigma_{27} < \sigma_{26}$ for the negatively signed squared lengths. Applying (3.28) to triangles 126 and 267 yields $\sigma_{26} = \sigma_{12}/\alpha_{126}^2 = \sigma_{12}/\alpha_{125}^2$ and $\sigma_{27} = \sigma_{26}/\alpha_{627}^2 = \sigma_{26}/\alpha_{425}^2$, where

$$\alpha_{ijk} \text{ refers to } \alpha \text{ of (3.13) for the angle bounded by edges } ij \text{ and } kj. \tag{3.29}$$

Both α_{125} and α_{425} can be determined using (3.13) in terms of the triangle squared lengths σ_t, σ_s on the boundary and σ in the interior. Therefore

$$\sigma_{27} = \frac{\sigma_{12}}{\alpha_{125}^2\alpha_{425}^2} \tag{3.30}$$

determines the light ray location in terms of σ_t, σ_s and $\sigma < 0$.

Similarly for a box with spacelike interior edges as in the middle of Figure 3.7,

$$\sigma_{37} = \sigma_{13}\alpha_{135}^2\alpha_{435}^2 \tag{3.31}$$

determines the light ray location in terms of σ_t, σ_s and $\sigma > 0$.

It is convenient to represent the light ray location using a dimensionless variable that grows linear with respect to length instead of squared length. We define the **dimensionless light ray location** as a function of the interior squared length σ by

$$r(\sigma) = \begin{cases} -\frac{\sqrt{\sigma_{47}}}{\sqrt{\sigma_{34}}} = \frac{\sqrt{\sigma_{37}}}{\sqrt{\sigma_{34}}} - 1 = \frac{\sqrt{\sigma_{13}\alpha_{135}^2\alpha_{435}^2}}{\sqrt{\sigma_{34}}} - 1, & \text{spacelike interior edge } \sigma > 0, \\ 0 & \text{lightlike interior edge } \sigma = 0, \\ \frac{\sqrt{\sigma_{47}}}{\sqrt{\sigma_{24}}} = 1 - \frac{\sqrt{\sigma_{27}}}{\sqrt{\sigma_{24}}} = 1 - \frac{\sqrt{\sigma_{12}\alpha_{125}^{-2}\alpha_{425}^{-2}}}{\sqrt{\sigma_{24}}}, & \text{timelike interior edge } \sigma < 0. \end{cases} \tag{3.32}$$

As the light ray location moves continuously from vertex 3 to 4 to 2, the value of r grows from -1 to 0 to 1 linearly so that, for instance, $r = -0.5$ when the light ray lands right in the middle between vertices 3 and 4.

3.3.2 Light ray fluctuations

Formula (3.32) allows one to compute the light ray location for a given configuration. When different configurations are summed over in a path integral, there are quantum fluctuations in the light ray location.

To characterize the quantum fluctuation of light ray location, we partition the possible values of r into N many equal size intervals $I_i, i = 1, 2, \cdots, N$. Path integrating over the spacetime configurations compatible with the set of light ray locations I_i yields the amplitude

$$A_i = \int_{\sigma: r(\sigma) \in I_i} d\sigma \; L[\sigma]C[\sigma]\mu[\sigma]e^{E[\sigma]}. \tag{3.33}$$

Given $r(\sigma)$, the values of σ corresponding to I_i can be solved numerically so that the integration domain is determined. From these amplitude, one could compute the relative probability for the light ray to land in the i-th interval (Figure 3.8)

$$p_i = \frac{|A_i|^2}{\sum_i |A_i|^2}, \tag{3.34}$$

The probability distribution over i informs us how much light ray fluctuation there is.

3.4 Light ray fluctuations in 2D

3.4.1 Fixed boundary size, varying coupling constants

In this section we fix the boundary squared lengths to $\sigma_s = 1, \sigma_t = -1$ and study the light ray fluctuation for different sets of coupling constants in 2D. Varying boundary squared lengths will be considered in the next section.

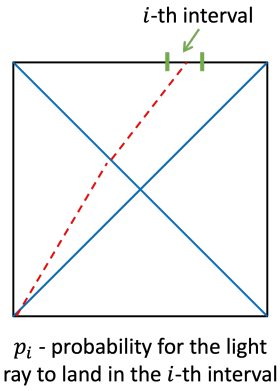

Figure 3.8: Partitioning the possible light ray locations into intervals and assigning relative probabilities according to the gravitational path integral.

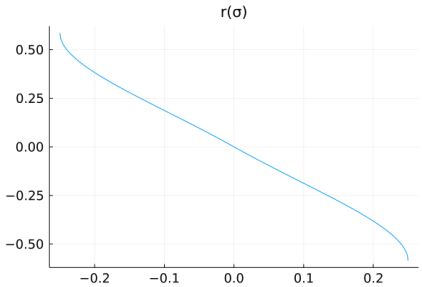

Figure 3.9: The light ray location $r(\sigma)$ in 2D as a function of the interior edge squared length σ when $\sigma_s = 1, \sigma_t = -1$.

The Lorentzian and lightcone constraints of Section 3.2.3 limit the interior edge squared length to $\sigma \in (\sigma_t/4, \sigma_s/4) = (-0.25, 0.25)$. Since the boundary values $\sigma = -0.25, 0.25$ are the branch point singularities for the path integral exponent, we take the integration domain for σ to be $[-b, b]$ where

$$b = 0.25 - \epsilon. \tag{3.35}$$

The choice of the cutoff value ϵ does not influence the probability distributions to be computed in any significant way, as long as it is kept small. Here we take $\epsilon = 10^{-8}$.

The light ray location $r(\sigma)$ as a function of the interior edge squared length is plotted in Figure 3.9 according to (3.32). Clearly the range of r is a proper subset of $[-1, 1]$, so not all locations on the boundary edges 34 and 24 are reachable by light rays emanating from vertex 1 in the current setting. We partition the *reachable* light ray locations into

$$N = 16 \tag{3.36}$$

equal size intervals, which are denoted I_i for $i = 1, 2, \cdots, 16$.

In the following, we compute the light location probabilities p_i for $i = 1, 2, \cdots, 16$ defined in (3.34) through numerical integration in the Julia programming language [84] using the QuadGK package based on the adaptive Gauss-Kronrod quadrature method.

The results are presented below in figure Figure 3.10 and onward. In all the figures, the coupling constants are displayed to 2 significant digits which explains the $\approx$ sign in the titles. Each bar chart shows a probability distribution of p_i kept to 4 decimal places with some fixed set of parameters. The path integral amplitudes as functions of σ are plotted below the bar charts. The measure factor $\mu[\sigma]$ is counted as a factor within the amplitude so that the integral is with respect to the plain Lebesgue measure in σ. From Figure 3.9 the light ray location is a decreasing function of σ, so the bar charts p_i are shown for $i = 16, \cdots, 1$ from left to right in order to match the increasing values of σ for the amplitude plots. This eases the comparison between the bar charts and the amplitude plots, and one observes that the places where the amplitude varies slowly corresponds well with the peaks of the probability distribution.

One non-vanishing parameter

As noted in Section 3.2.1, the Gauss-Bonnet theorem implies that the Einstein-Hilbert term is a constant in the present model with fixed topology. This constant term drops out as a common factor for the numerator and denominator in the definition (3.34) for p_i. Therefore there are only three non-trivial free parameters m, λ, a. We first consider the cases where only one of the three parameters is non-zero.

When a parameter $x = \lambda$ or a is non-vanishing, it can be both positive and negative. However, p_i for x equals p_i for $-x$, because for the amplitudes are complex conjugates for all configurations. For this reason we only display the data for positive λ and a in Figure 3.10 and Figure 3.11.

When only $\lambda > 0$ the results are shown in Figure 3.10. For large $\lambda \gtrsim 1000$ light ray fluctuation is small as p_i is sharply peaked. As λ gets smaller the amount of fluctuation gets larger. A similar conclusion holds when only $a > 0$ (Figure 3.11), namely light ray fluctuation is small and large respectively when a is large and small. The reason is clear from the amplitude plots below the

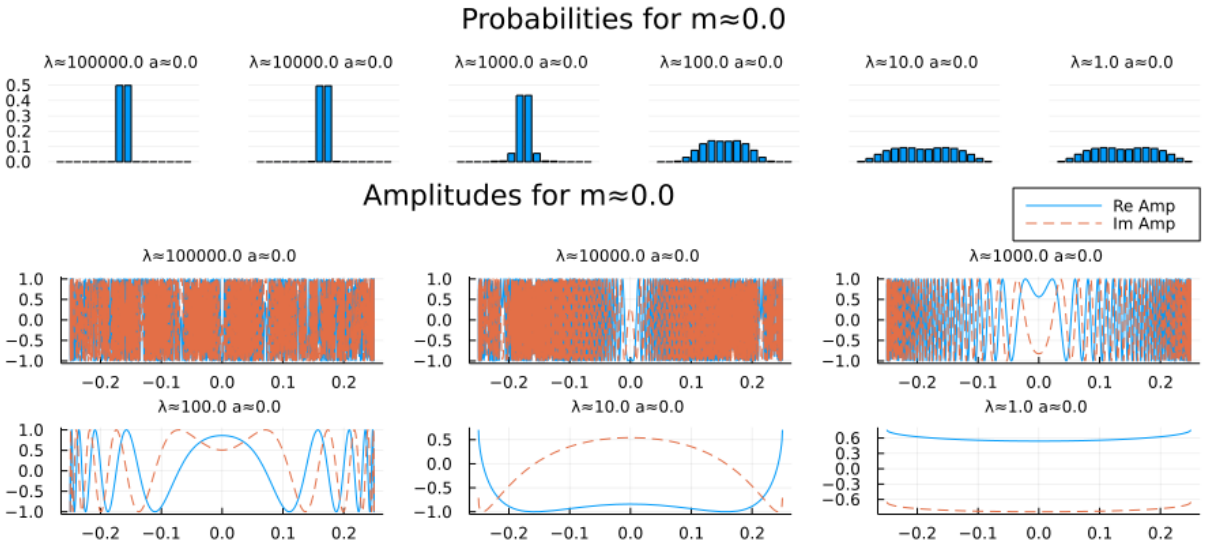

Figure 3.10: Probability and amplitude distributions for a family of λ with $m = 0, a = 0$. By Figure 3.9 the light ray location is a decreasing function of σ. Therefore in this and the following figures of $2D$, the probabilities p_i of (3.34) are plotted for $i = 16, \cdots, 1$ from left to right in order to match the increasing values of σ for the amplitude plots.

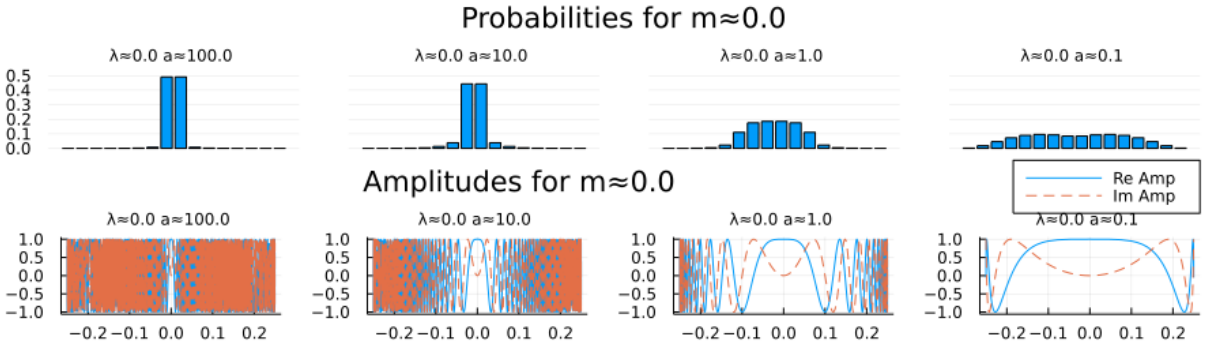

Figure 3.11: Probability and amplitude distributions for a family of a with $m = 0, \lambda = 0$.

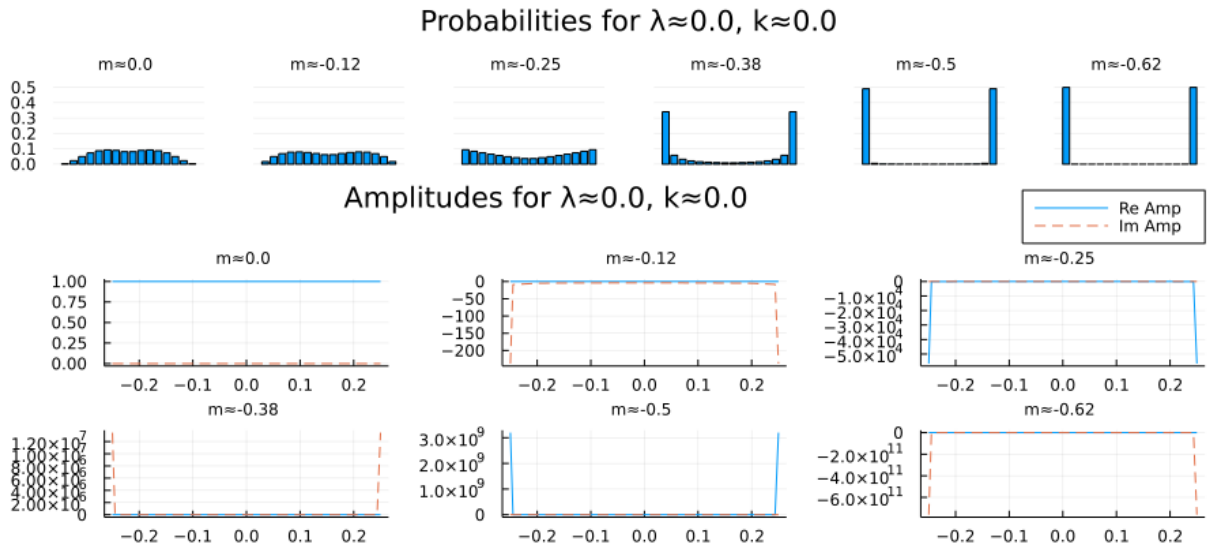

Figure 3.12: Probability and amplitude distributions for a family of m with $\lambda = 0, a = 0$.

probabilities. A large parameter makes the phase of the integral amplitude change faster so that probabilities are suppressed except for special regions where the phase is nearly stationary.

These fit the common intuition that as $\hbar$ gets smaller quantum fluctuations become smaller. The path integral exponent E of (3.6) is related to the action S by $E = \frac{i}{\hbar}S$, so the coupling constants λ and a scale inverse-proportionally with $\hbar$. Therefore a smaller $\hbar$ means a large absolute value for λ and a, and we saw that these yield smaller light ray fluctuations.

Next we consider $m < 0$ with the results shown in Figure 3.12. The p_i values start distributed evenly for $m = 0$, and gets pushed towards the two sides as m decreases. This trend is easy understand. In the definition (3.17) $\mathbb{V}_s$ is always negative so for the probabilities only $|\mathbb{V}_s|$ matters. The smaller the negative exponent m is, the more it favours configurations with small $|\mathbb{V}_s|$, which in the current model means larger values of $|\sigma|$ according to (3.8). Therefore for a very small negative m the light ray location is concentrated around the two sides of the plots.

In summary if x is the only non-vanishing parameter among the three parameters, large $|x|$ suppresses light ray fluctuations while small $|x|$ enhances light ray fluctuations.

Multiple non-vanishing parameters

When multiple parameters are non-vanishing, one could expect that the qualitative features of the probability distribution follows that of the dominating parameter. This indeed holds in the cases studied below.

For reasons that will become clear in Section 3.4.2, we organize the set of parameters according to the product λa. We consider $m = 0$ and $m = -0.25$ (the value for the Dewitt measure [18]) but not smaller values as they only serve to push the probability distribution towards the two sides like in Figure 3.12.

The results for $\lambda a = 10^4$ are shown in Figure 3.13 and Figure 3.14. Light ray fluctuation is small for all the 6 families of parameters shown. Here with $\lambda a = 10^4$, there is always at least one

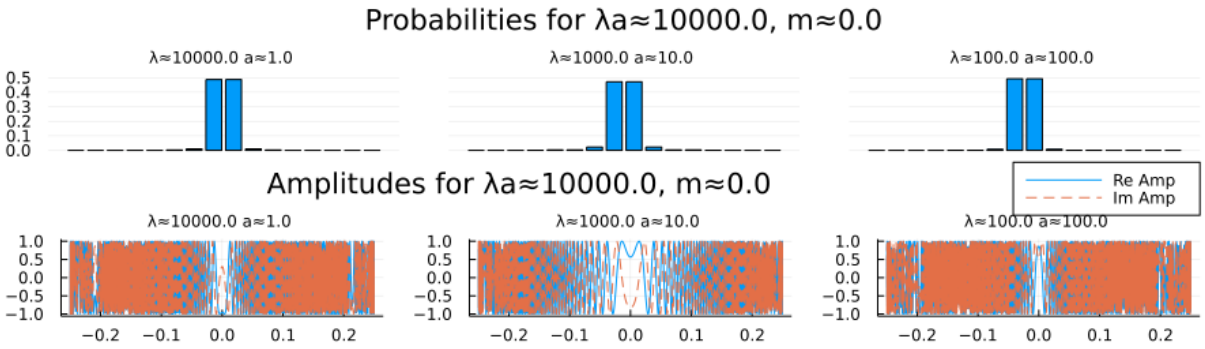

Figure 3.13: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 10^4, m = 0$.

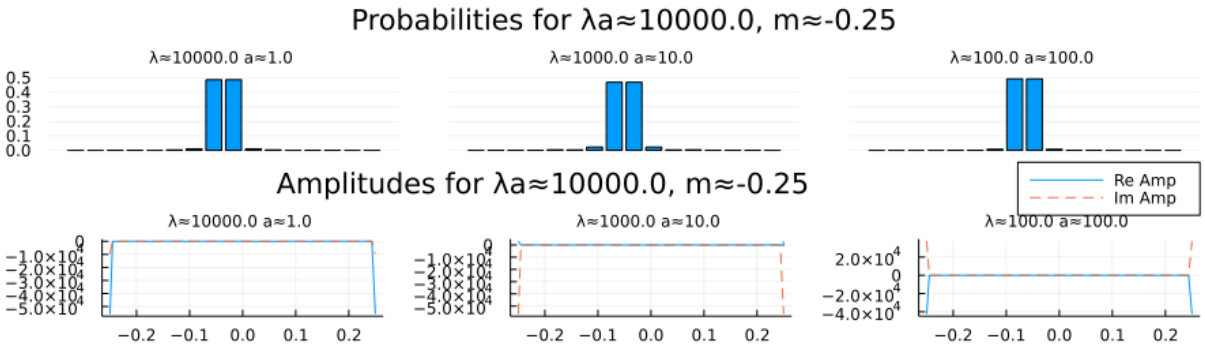

Figure 3.14: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 10^4, m = -0.25$.

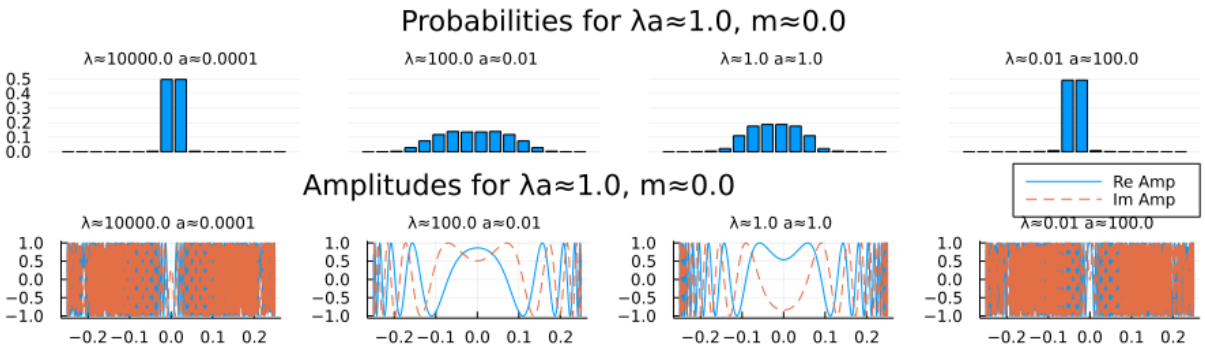

Figure 3.15: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 1, m = 0$.

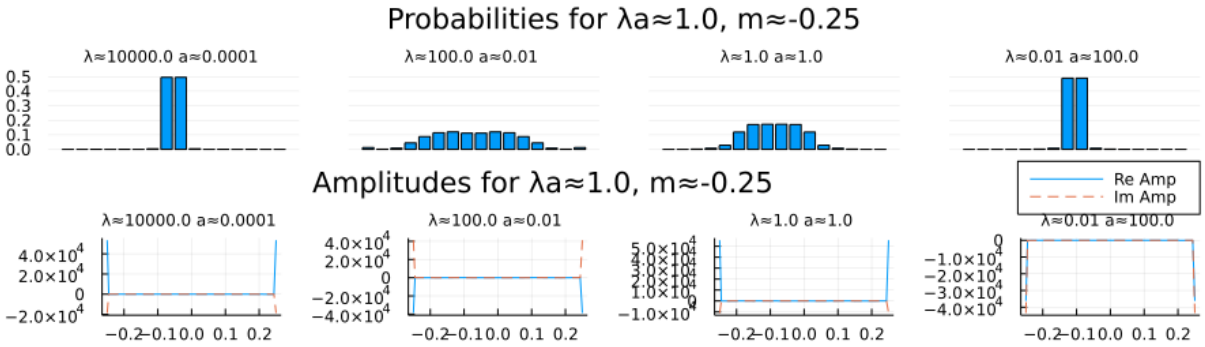

Figure 3.16: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 1, m = -0.25$.

of λ and a that is large to yield fast oscillation of the phase and suppress the probabilities away from the center region.

The results for $\lambda a = 1$ is shown in Figure 3.15 and Figure 3.16. For this smaller value of λa as well, the probabilities are concentrated around the center when $\lambda \gtrsim 1000$ or $a \gtrsim 100$. Different from the case of $\lambda a = 10^4$, there are now families of parameters in which both λ and a are small (e.g., $\lambda = 1, a = 1$) so that light ray fluctuation is larger.

As shown in Figure 3.17 and Figure 3.18, the results for a yet smaller positive value $\lambda a = 10^{-6}$ exhibits the same qualitative features, which we expect to hold generically for small positive values of λa.

With two or more non-vanishing parameters it makes a difference to allow for negative values of λa. The results for $\lambda a = -10^4$ is shown in Figure 3.19 and Figure 3.20. In comparison with $\lambda a = 10^4$, the only major difference is that for $\lambda = 1000$, light ray fluctuation is less suppressed. A comparison of the amplitude plots show that the negative $a = -10$ induces a wider region of slowly changing phases, which explains the difference in probabilities.

We also studied other negative values of λa opposite to the positive ones studied above. The plots turned out to be similar to the corresponding positive values so are not shown.

In summary, when $\lambda, a \neq 0$, the presence of a large $|\lambda|$ or $|a|$ suppresses light ray fluctuations while their absence enhances light ray fluctuations. In special cases, changing a parameter with

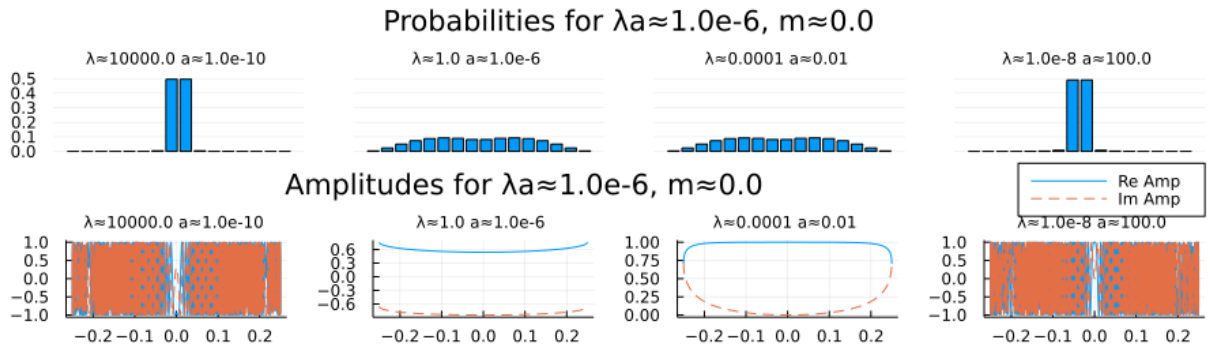

Figure 3.17: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 10^{-6}, m = 0$.

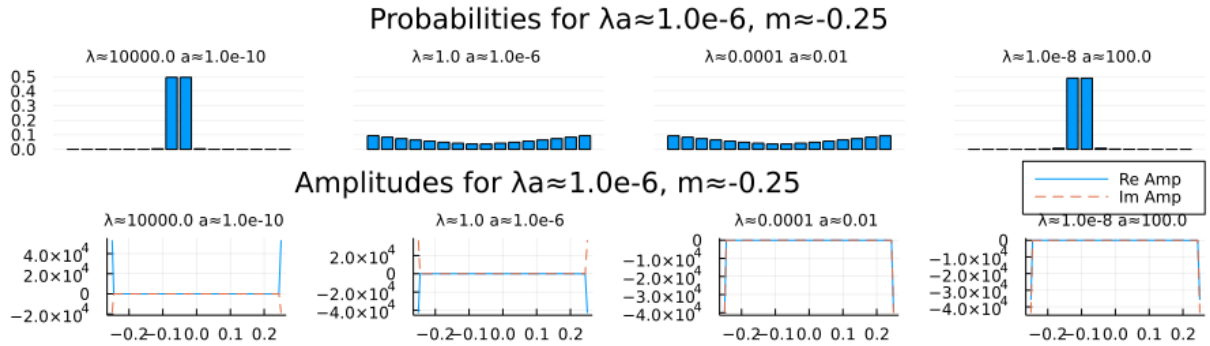

Figure 3.18: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = 10^{-6}, m = -0.25$.

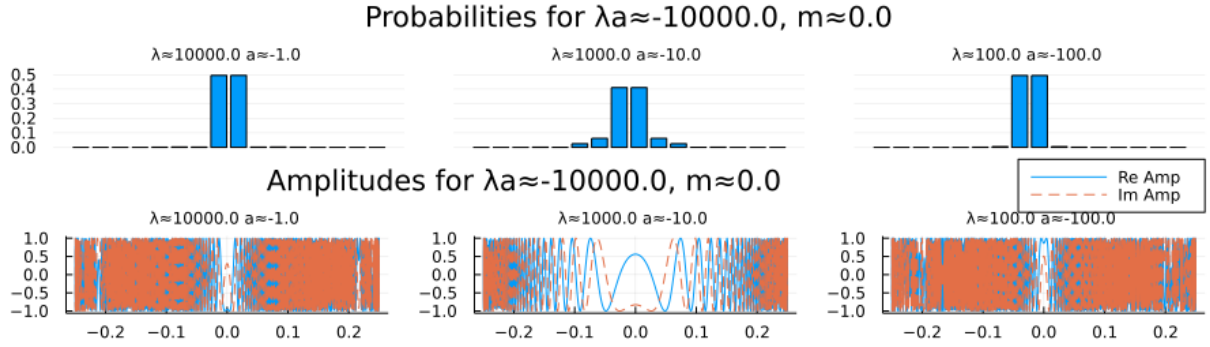

Figure 3.19: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = -10^4, m = 0$.

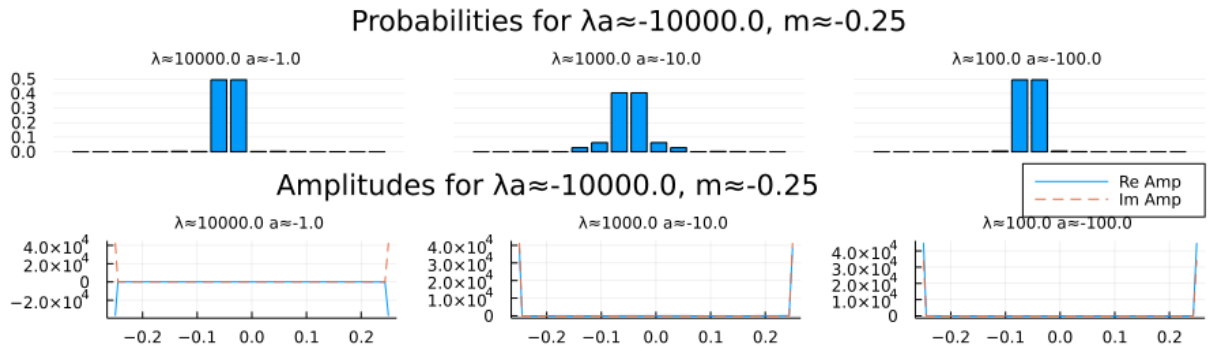

Figure 3.20: Probability and amplitude distributions for a family of (λ, a) with $\lambda a = -10^4, m = -0.25$.

smaller magnitude (e.g., $\lambda = 1000, a = 10$ to $\lambda = 1000, a = -10$) can mildly affect the amount of light ray fluctuation due to cancellations between terms in the exponent.

3.4.2 Fixed coupling constants, varying boundary sizes

To study how light ray fluctuations are affected by the size of the region specified by the boundary edge lengths, we can exploit the scaling identity of Section 3.2.4. In 2D (3.25) implies that

$$Z[l^2\sigma_B, \lambda, k, a, m] = l^{2N_e+4mN_t}\, Z[\sigma_B, l^2\lambda, k, l^{-2}a, m]. \tag{3.37}$$

Since the light ray location r is a dimensionless quantity, one could easily check that the equation holds when Z is replaced by A_i of (3.33). Consequently

$$p_i[l^2\sigma_B, \lambda, k, a, m] = p_i[\sigma_B, l^2\lambda, k, l^{-2}a, m]. \tag{3.38}$$

This says that rescaling the boundary condition (LHS) is equiavlent to rescaling the parameters λ and a while keeping λa fixed. Therefore the previous results for fixed λa also inform us about how light ray fluctuations depend on the size of the region. For instance, suppose we fix the parameters to be $\lambda = 10^4, a = 1, m = 0$. Then for $\sigma_s = -\sigma_t = 1$, p_i are as in the first plot in Figure 3.13. The next plots are for the boundary squared lengths shrinked by factors of $l^2 = 0.1$.

All the previous figures for fixed λa can then be read from left to right as a progressive shrinking of the size of the boundary for fixed parameters. We see that light ray fluctuation is suppressed for both large and small sized boundaries.

3.4.3 Time-space duality

A notable feature of the data for p_i is that the values of p_i are completely symmetrical in the sense that $p_i = p_{N-i}$ for $i = 1, 2, \cdots, N/2$. This is not a coincidence but follows from some symmetry considerations.

Given a path integral configuration σ, consider the map $\sigma \mapsto -\sigma$ that negates all the squared lengths including the boundary ones. Physically, this map exchanges spacelike intervals and

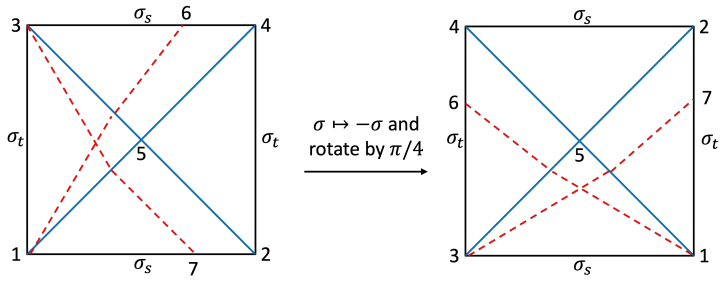

Figure 3.21: Applying $\sigma \mapsto -\sigma$ followed by a $\pi/4$ rotation to a box configuration obeying $\sigma_t = -\sigma_s$ yields a box configuration obeying the same boundary condition. The light ray location r at 6 is changed to 7, which by the symmetry of the model is located at $-r$.

timelike intervals. It was noted that this map is a symmetry for 2D Lorentzian quantum gravity with the cosmological constant term and the Einstein-Hilbert term [5].

Here we note that the symmetry holds more generally for any theory with additional even powers of the Ricci scalar R in the action. According to Section 3.2.1, the simplicial analogue of R is $-\delta_v/A_v$. From the definitions of δ_v and A_v, $\sigma \mapsto -\sigma$ maps δ_v to $-\delta_v$ and A_v to A_v. Hence $-\delta_v/A_v$ is mapped to δ_v/A_v, and any even power of $-\delta_v/A_v$ is left invariant. Since the map $g \mapsto -g$ on the metric field g also takes R to $-R$, the symmetry is also expected to hold for additional theories of 2D quantum gravity beyond simplicial quantum gravity.

While the map $\sigma \mapsto -\sigma$ exchanges timelike and spacelike intervals, it preserves lightlike intervals. Applied to a box model configuration it turns the timelike boundaries to spacelike ones and *vice versa*, while preserving all light ray paths. For the models studied the boundary condition obeys $\sigma_t = -\sigma_s$, so a rotation of the box by $\pi/4$ in either direction yields back a configuration obeying the same boundary condition. On the other hand the light ray location is changed from r to $-r$ Figure 3.21. Since the map $\sigma \mapsto -\sigma$ preserves path integral amplitudes, we see that each box configuration with light ray location r corresponds to another configuration of the same amplitude with light ray location $-r$. In the 2D symmetry-reduced model $r \in I_i$ if and only if $-r \in I_{N-i}$, so $A_i = A_{N-i}$ and $p_i = p_{N-i}$.

3.5 Light ray fluctuations in 3D

3.5.1 Fixed boundary size, varying coupling constants

The study in 3D is parallel to the case of 2D. As illustrated in Figure 3.6, the light ray travels in the diagonal plane. Again we fix the boundary squared lengths to $\sigma_s = 1, \sigma_t = -1$. The Lorentzian and lightcone constraints limit the interior edge squared length to $\sigma \in ((\sigma_s + \sigma_t)/4, \sigma_s/2) = (0.0, 0.5)$. In particular, four light rays emanate from the interior vertex in the diagonal plane if and only if two lightcones are attached to the interior vertex in the 3D box, so this condition is enough to ensure the lightcone constraint is obeyed. Like in 2D, the integration domain is cutoff with $\epsilon = 10^{-8}$ at

$$(\epsilon, 0.5 - \epsilon) \tag{3.39}$$

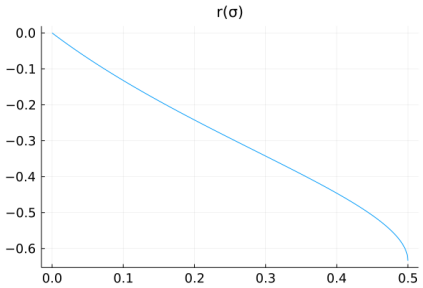

Figure 3.22: The light ray location $r(\sigma)$ in 3D as a function of the interior edge squared length σ when $\sigma_s = 1, \sigma_t = -1$.

to avoid branch point singularities in the domain. The choice of the cutoff value ϵ does not influence the probability distributions to be computed in any significant way.

In computing the probabilities through numerical integration we need to impose an additional cutoff around $\sigma = \sigma_s/4$. At this point some angles evaluated according to (3.12) diverge, but the results should be finite since the divergences of different angles cancel in path integral exponent. However to avoid numerical inaccuracies due to the difference of large numbers, a cutoff is imposed. Where $|\sigma - \sigma_s/4| < \epsilon' = 10^{-7}$ in the integration domain, the integrand is evaluated at $\sigma = \sigma_s/4 \pm \epsilon'$ for $\sigma >, < \sigma_s/4$. This value of ϵ' is picked to be small without leading to numerical infinities.

The light ray location $r(\sigma)$ as a function of the interior edge squared length is plotted in Figure 3.22 according to (3.32) when $\sigma_{12} = \sigma_{34} = 2\sigma_s = 2$ as illustrated in Figure 3.6. The reachable light ray locations are again partitioned into

$$N = 16 \tag{3.40}$$

equal size intervals, which are denoted I_i for $i = 1, 2, \cdots, 16$.

In 3D the parameters are λ, k, m for the cosmological constant term, the Einstein-Hilbert term, and the measure term. In figures such as Figure 3.23 the probability distributions of p_i to 4 decimal places are plotted in the bar charts for $i = 16, \cdots, 1$ from left to right to match the increasing values of σ for the amplitude plots. The range of probability plots is limited to $[0.0, 0.5]$ to make the low probability bars more visible.

One non-vanishing parameter

We first consider the cases where only one of the three parameters is non-zero. We only display the results for positive λ and k in Figure 3.23 and Figure 3.24, because as in the 2D case p_i for the positive and negative parameters are equal.

For large $\lambda \gtrsim 1000$ or $k \gtrsim 1000$, p_i is sharply peaked indicating small light ray fluctuations. As λ or k gets smaller the fluctuation gets larger. The reason is clear from the amplitude plots below the probabilities. As in the 2D case, the large parameter makes the phase of the integral amplitude change faster so that probabilities are suppressed except for special regions where the phase is nearly stationary. As in the 2D case, these fit the common intuition that as $\hbar$ gets smaller quantum fluctuations become smaller. The path integral exponent E of (3.14) is related to the action S by $E = \frac{i}{\hbar}S$, so the coupling constants λ and k scale inverse-proportionally with

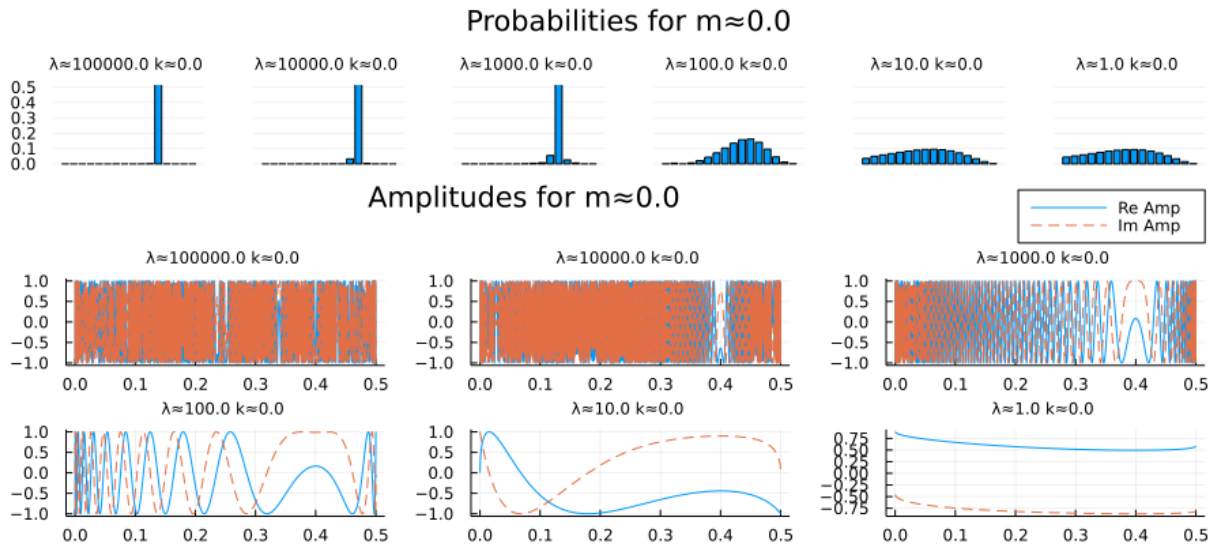

Figure 3.23: Probability and amplitude distributions for a family of λ with $m = 0, k = 0$. By Figure 3.22 the light ray location is a decreasing function of σ. Therefore in this and the following figures of $3D$, the probabilities p_i of (3.34) are plotted for $i = 16, \cdots, 1$ from left to right in order to match the increasing values of σ for the amplitude plots.

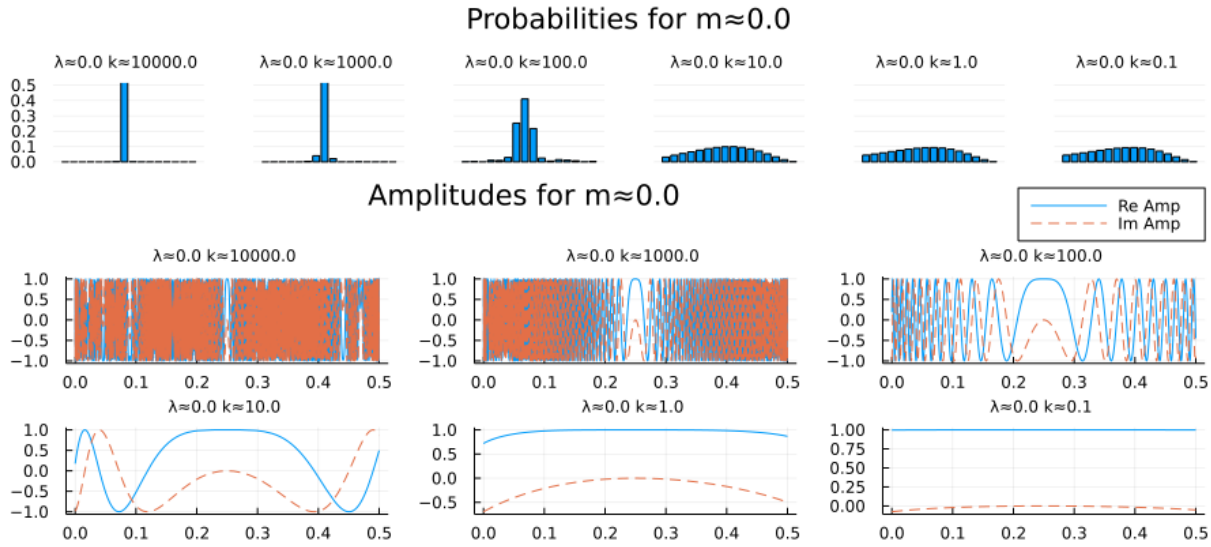

Figure 3.24: Probability and amplitude distributions for a family of a with $m = 0, \lambda = 0$.

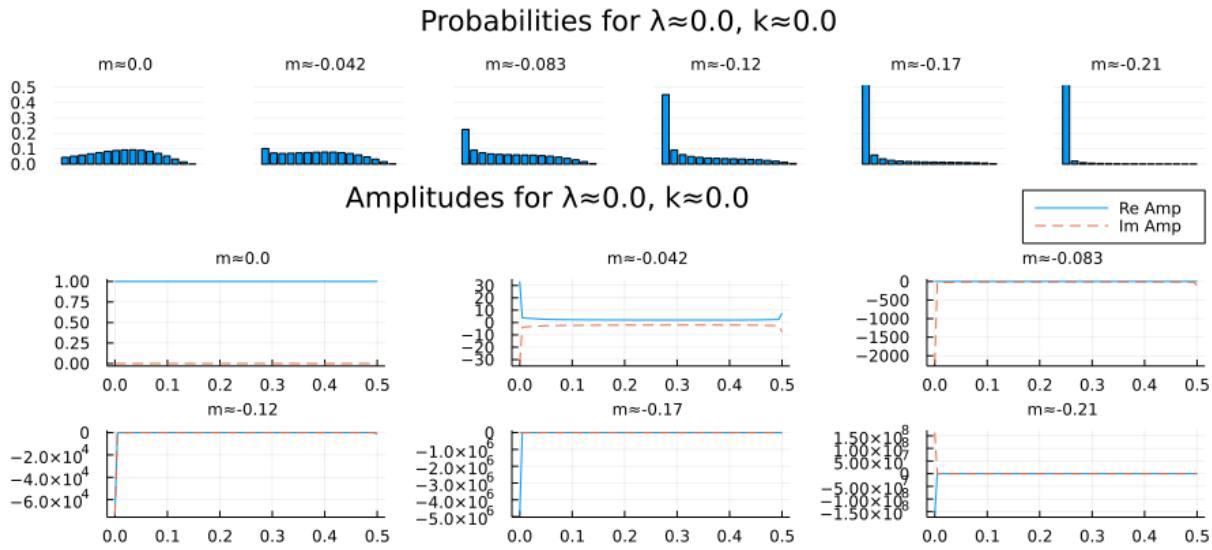

Figure 3.25: Probability and amplitude distributions for a family of m with $\lambda = 0, k = 0$.

$\hbar$. Therefore a smaller $\hbar$ means a larger absolute value for λ and a, and we see that these yield smaller light ray fluctuations.

The results for $m < 0$ are shown in Figure 3.25. The p_i values start distributed evenly for $m = 0$, and gets pushed towards the left side as m decreases. Like in 2D, the smaller the negative parameter m is, the more it favours configurations with small $|\mathbb{V}_s|$. In 3D this favours smaller values of σ according to (3.15) and explains why the light ray location is concentrated around the left side of the plots.

In summary if x is the only non-vanishing parameter among the three parameters, large $|x|$ suppresses light ray fluctuations while small $|x|$ enhances light ray fluctuations.

Multiple non-vanishing parameters

In 3D the scaling analysis requires fixing k^3/λ (Section 3.5.2), so the parameters are organized according to the value of k^3/λ. We consider $m = 0$ and $m = -1/12$ (the value for the Dewitt measure [18]) but not smaller values as they only serve to push the probability distribution towards the side like in Figure 3.25.

The results for $k^3/\lambda = 10^6$ are shown in Figure 3.26 and Figure 3.27. For fixed $k^3/\lambda > 0$, λ is large when k is large. The plots show that light ray fluctuation is large for large values of λ and k, and small for small values of λ and k. This is due to larger values of the parameters inducing fast oscillations for the phase away from the nearly stationary phase region. The results for other values of k^3/λ including negative ones are shown in Figure 3.28 to Figure 3.33 and exhibit no qualitative difference.

The situation is contrasted with 2D where as λ or a is decreased the light ray fluctuation starts large and ends large. Clearly this is because in 2D λa is dimensionless and fixed so λ and a are inversely proportional, whereas in 3D $k^3/\lambda > 0$ is dimensionless and fixed so λ and k^3 are proportional.

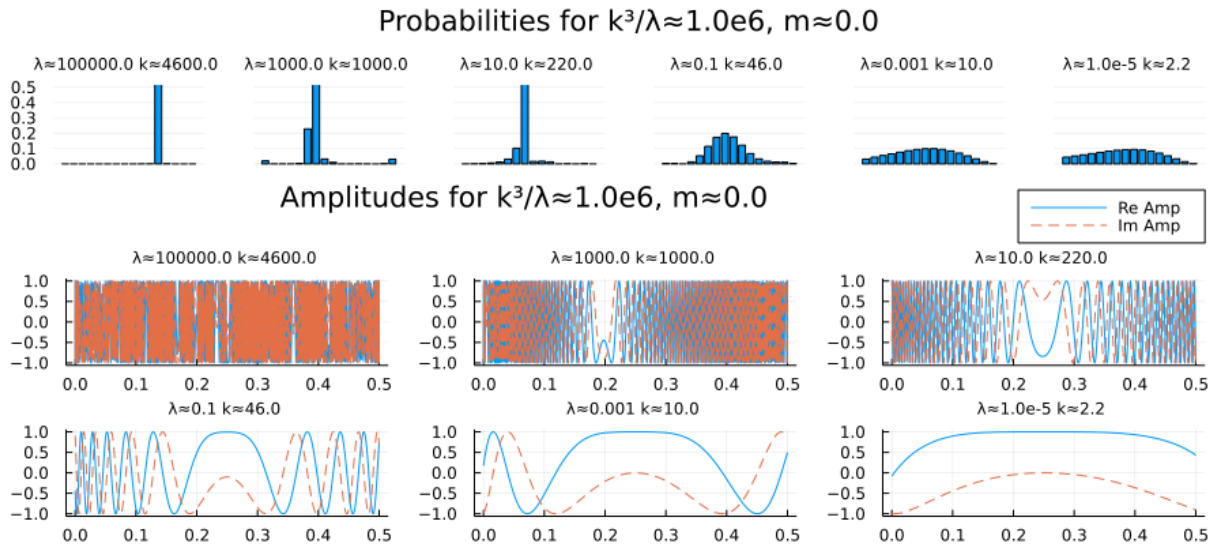

Figure 3.26: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = 10^6, m = 0$.

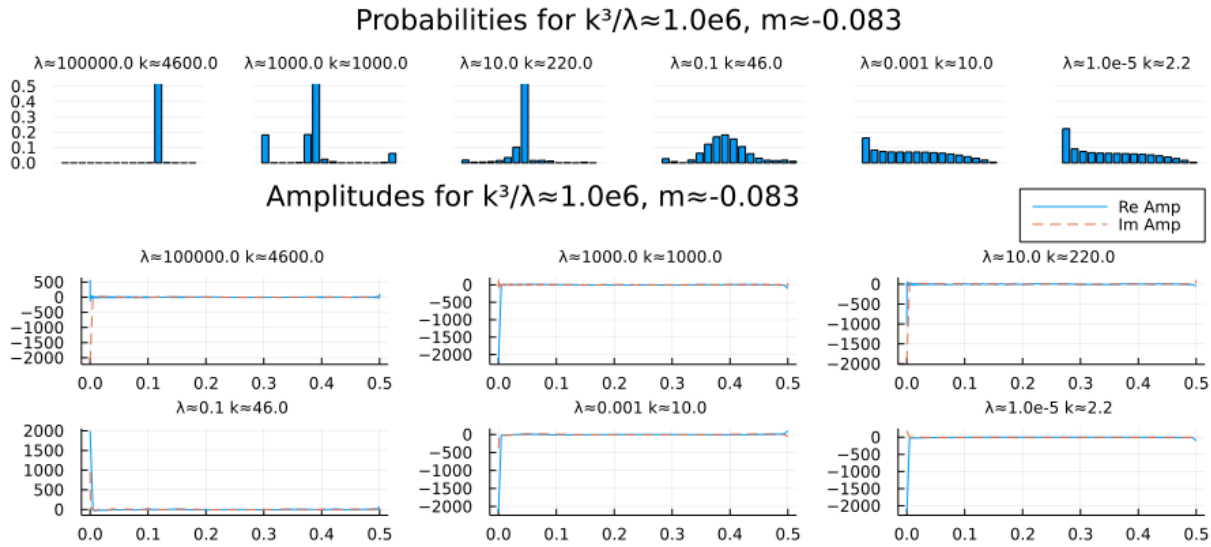

Figure 3.27: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = 10^6, m \approx -1/12$.

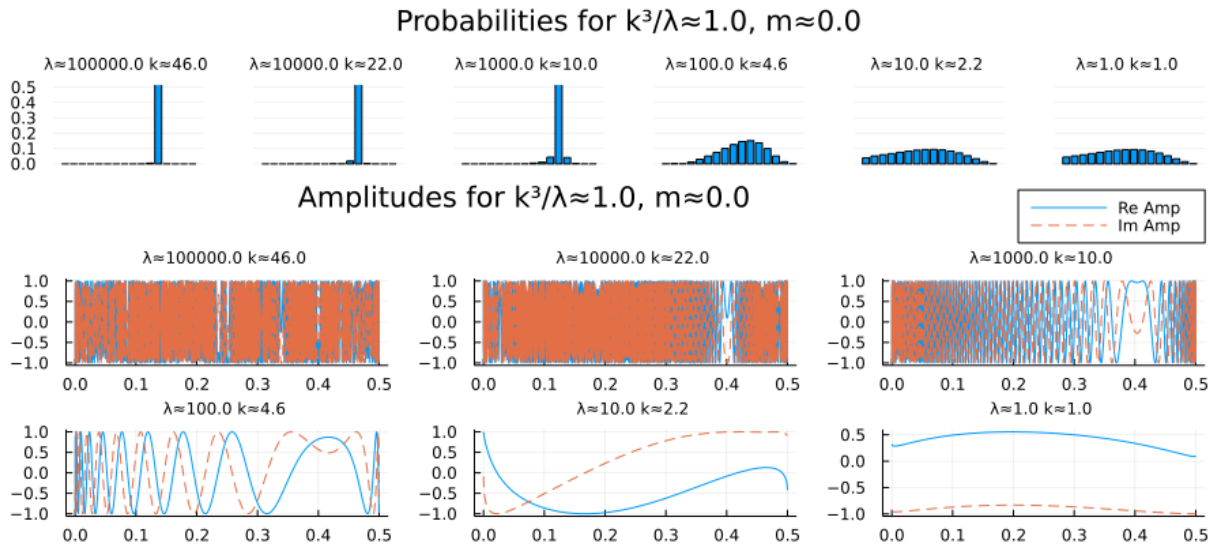

Figure 3.28: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = 1, m = 0$.

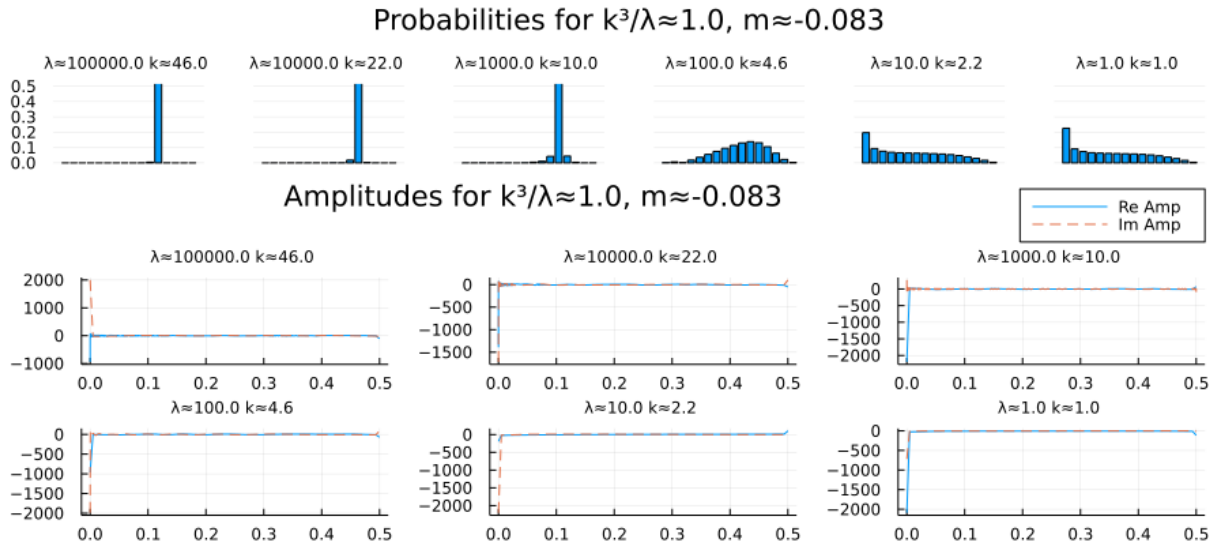

Figure 3.29: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = 1, m \approx -1/12$.

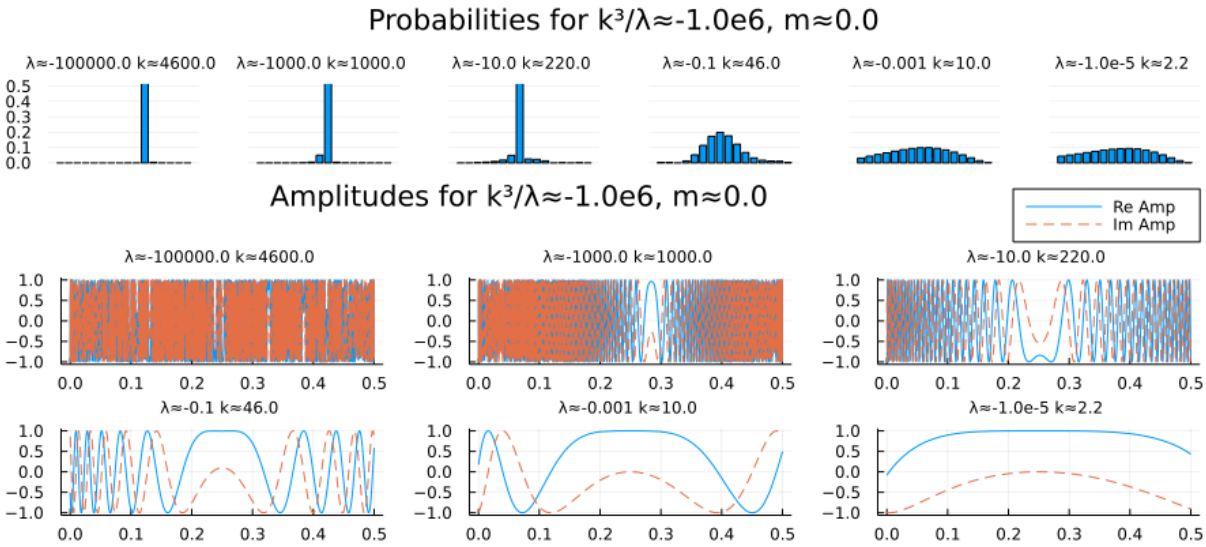

Figure 3.30: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = -10^6, m = 0$.

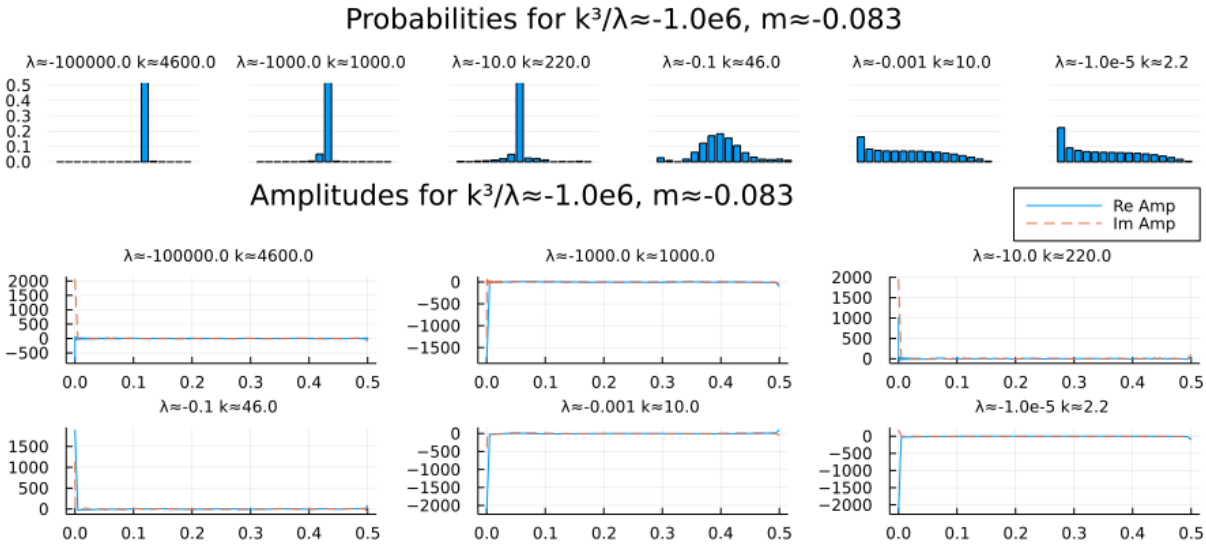

Figure 3.31: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = -10^6, m \approx -1/12$.

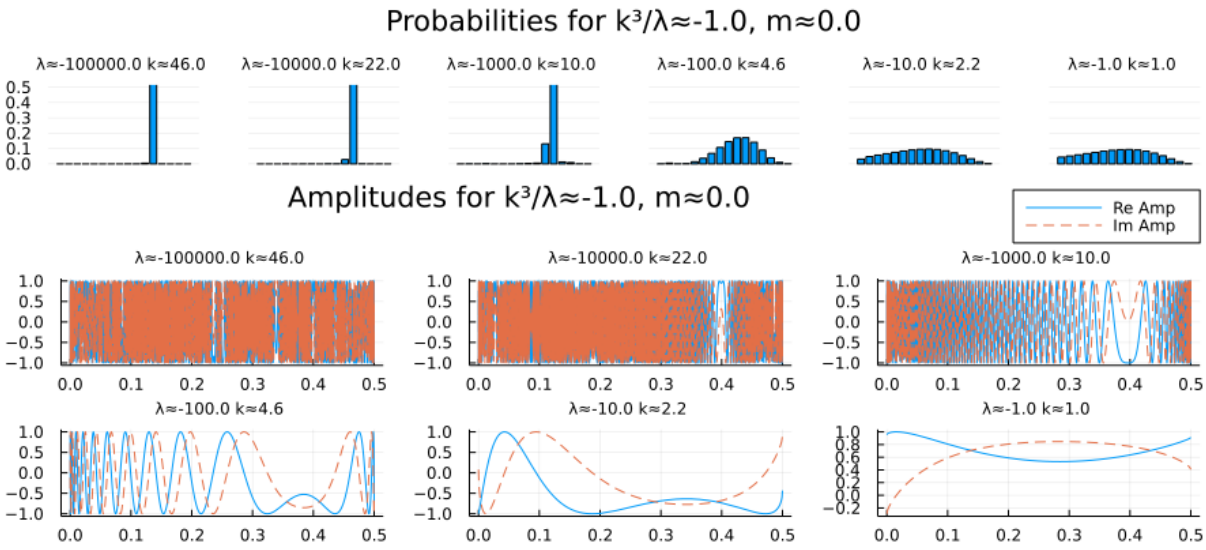

Figure 3.32: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = -1, m = 0$.

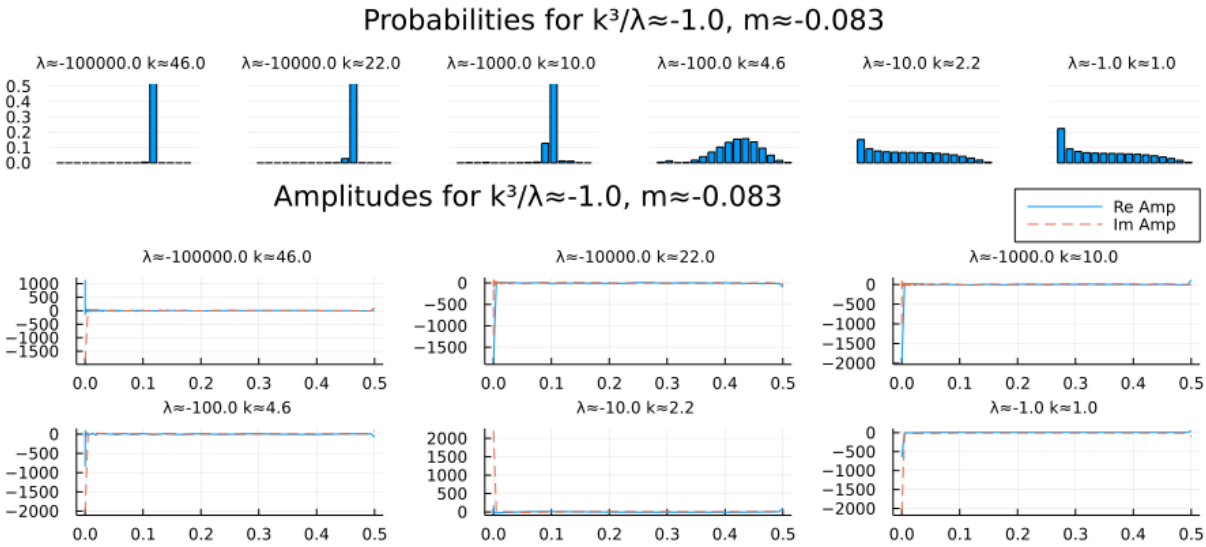

Figure 3.33: Probability and amplitude distributions for a family of (λ, k) with $k^3/\lambda = -1, m \approx -1/12$.

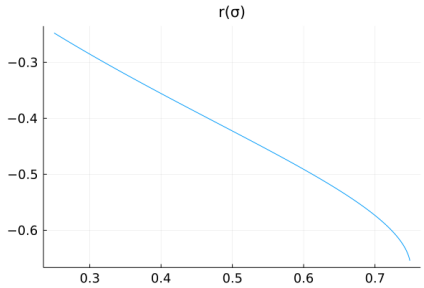

Figure 3.34: The light ray location $r(\sigma)$ in 4D as a function of the interior edge squared length σ when $\sigma_s = 1, \sigma_t = -1$.

3.5.2 Fixed coupling constants, varying boundary sizes

In 3D the scaling identity (3.25) implies that

$$Z[l^2\sigma_B, \lambda, k, m] = l^{2N_e+4mN_t}\ Z[\sigma_B, l^3\lambda, lk, m], \tag{3.41}$$

$$p_i[l^2\sigma_B, \lambda, k, a, m] = p_i[\sigma_B, l^3\lambda, lk, m]. \tag{3.42}$$

Therefore all the previous figures for fixed k^3/λ can then be read from left to right as a progressive shrinking of the size of the boundary for fixed parameters. We see that light ray fluctuation increases as the boundary shrinks.

3.6 Light ray fluctuations in 4D

3.6.1 Fixed boundary size, varying coupling constants

The setting of 4D is similar to the previous ones. Again we fix the boundary squared lengths to $\sigma_s = 1, \sigma_t = -1$ and the light ray travels in the diagonal plane as illustrated in Figure 3.6. The Lorentzian and lightcone constraints limit the interior edge squared length to $\sigma \in (\sigma_s/2 + \sigma_t/4, 3\sigma_s/4 = (0.25, 0.75)$, and the integration domain is cutoff with $\epsilon = 10^{-8}$ at

$$(0.25 + \epsilon, 0.75 - \epsilon). \tag{3.43}$$

In numerical integration, an additional cutoff is imposed to evaluate the integrand at $\sigma = \sigma_s/2 \pm \epsilon'$ for $\sigma >, < \sigma_s/2$ when $|\sigma - \sigma_s/4| < \epsilon' = 10^{-7}$. The light ray location $r(\sigma)$ as a function of the interior edge squared length is plotted in Figure 3.34 according to (3.32) when $\sigma_{12} = \sigma_{34} = 3\sigma_s = 3$. The reachable light ray locations are again partitioned into $N = 16$ equal size intervals, which are denoted I_i for $i = 1, 2, \cdots, 16$. The parameters are λ, k, m for the cosmological constant term, the Einstein-Hilbert term, and the measure term.

The probabilities and amplitudes are plotted in figures Figure 3.35 to Figure 3.40. Again, the probabilities p_i are kept to 4 decimal places, plotted in bars with the sequence $i = 16, \cdots, 1$ to match the increasing values of σ for the amplitude plots, and the plot range is limited to $[0.0, 0.5]$ to make the low probability bars more visible. The results are very similar to 3D, so we just briefly summarize them.

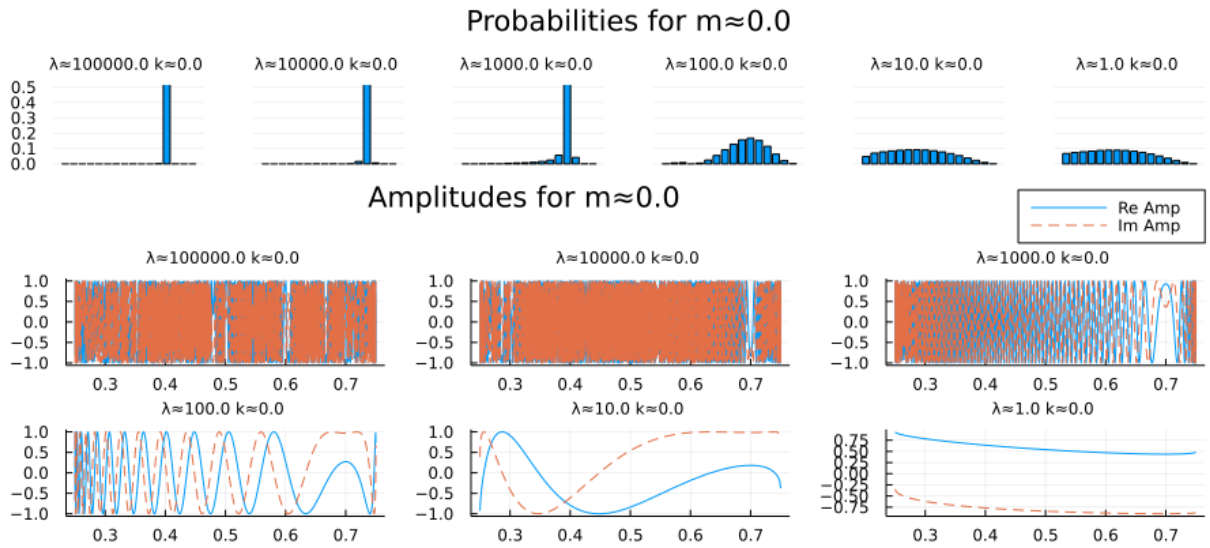

Figure 3.35: Probability and amplitude distributions for a family of λ with $m = 0, a = 0$. By Figure 3.34 the light ray location is a decreasing function of σ. Therefore in this and the following figures of $4D$, the probabilities p_i of (3.34) are plotted for $i = 16, \cdots, 1$ from left to right in order to match the increasing values of σ for the amplitude plots.

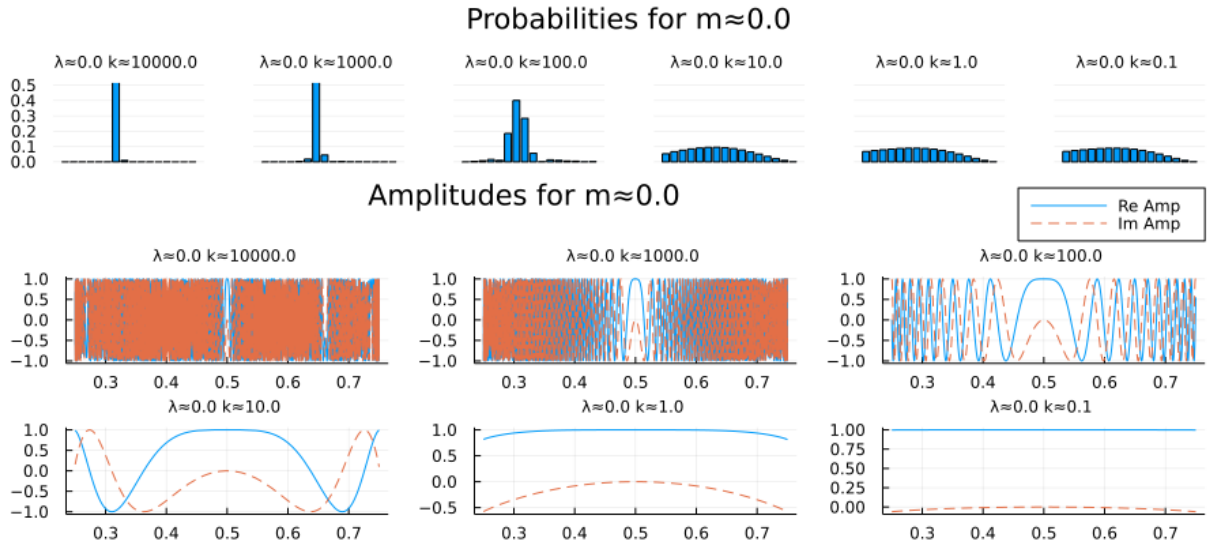

Figure 3.36: Probability and amplitude distributions for a family of a with $m = 0, \lambda = 0$.

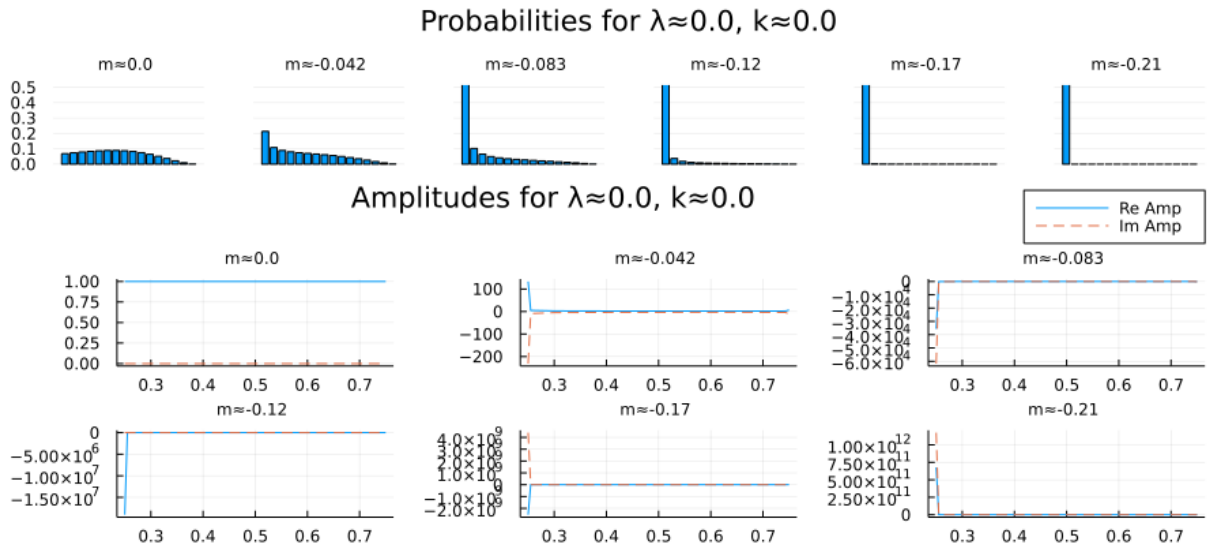

Figure 3.37: Probability and amplitude distributions for a family of m with $\lambda = 0, a = 0$.

One non-vanishing parameter

The results when only one parameter x of the three parameters is non-zero are plotted in Figure 3.35 to Figure 3.37. Similar to the 3D cases, large $|x|$ suppresses light ray fluctuations while small $|x|$ enhances light ray fluctuations. These fit the common intuition that as $\hbar$ gets smaller quantum fluctuations become smaller, since according to (3.14) and $E = \frac{i}{\hbar}S$, $|x|$ scale inverse-proportionally with $\hbar$, and we see that larger $|x|$ yield smaller light ray fluctuations.

Multiple non-vanishing parameters

The results for some families of non-vanishing (λ, k) are shown in Figure 3.38 to Figure 3.41. The parameters are organized according to the value of k^2/λ, which is fixed in the scaling analysis in 4D (Section 3.6.2). We consider $m = 0$, the value for the Dewitt measure [18], when smaller values only serve to push the probability distribution towards the side like in Figure 3.37. As in 3D light ray fluctuation is large for large values of λ and k, and small for small values of λ and k.

3.6.2 Fixed coupling constants, varying boundary sizes

In 4D the scaling identity (3.25) implies that

$$Z[l^2\sigma_B, \lambda, k, m] = l^{2N_e+4mN_t}\, Z[\sigma_B, l^4\lambda, l^2k, m], \tag{3.44}$$

$$p_i[l^2\sigma_B, \lambda, k, a, m] = p_i[\sigma_B, l^4\lambda, l^2k, m]. \tag{3.45}$$

Therefore all the previous figures for fixed k^2/λ can then be read from left to right as a progressive shrinking of the size of the boundary for fixed parameters. As in 3D, light ray fluctuation increases as the boundary shrinks.

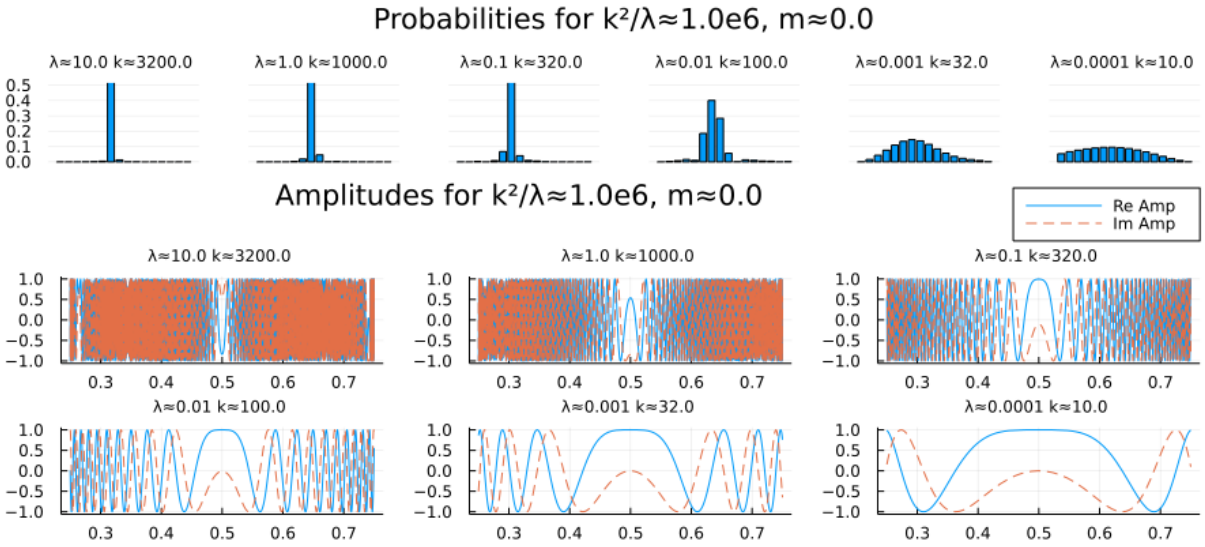

Figure 3.38: Probability and amplitude distributions for a family of (λ, k) with $k^2/\lambda = 10^6, m = 0$.

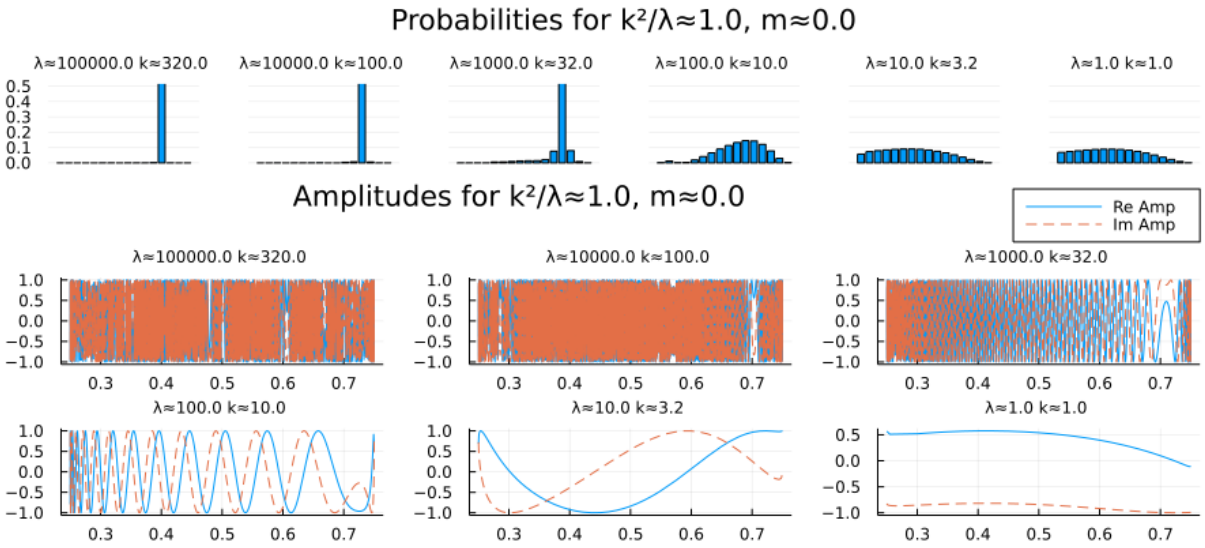

Figure 3.39: Probability and amplitude distributions for a family of (λ, k) with $k^2/\lambda = 1, m = 0$.

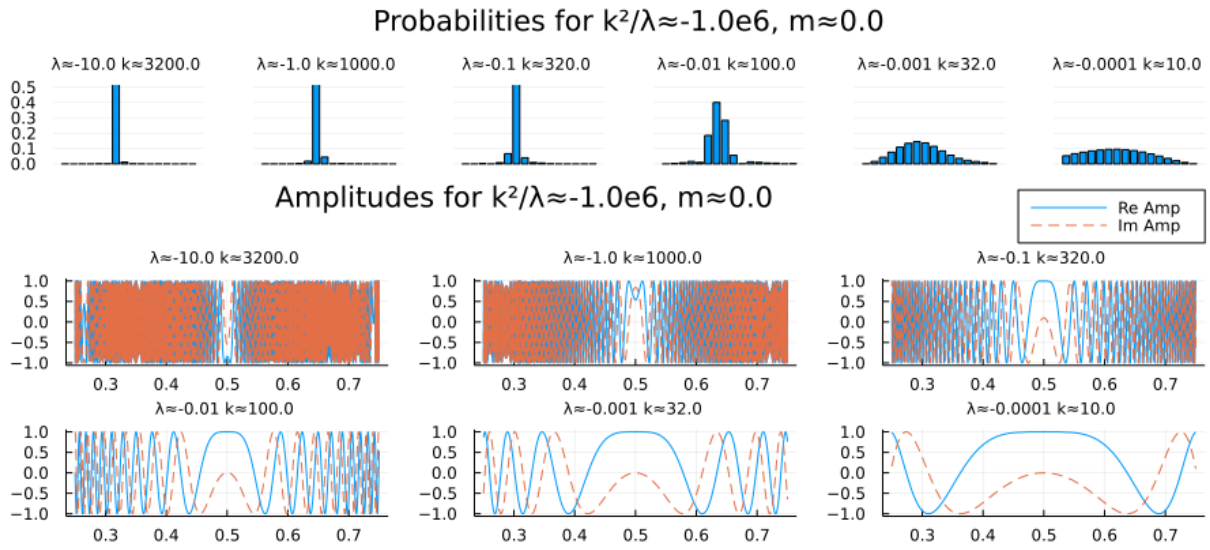

Figure 3.40: Probability and amplitude distributions for a family of (λ, k) with $k^2/\lambda = -10^6, m = 0$.

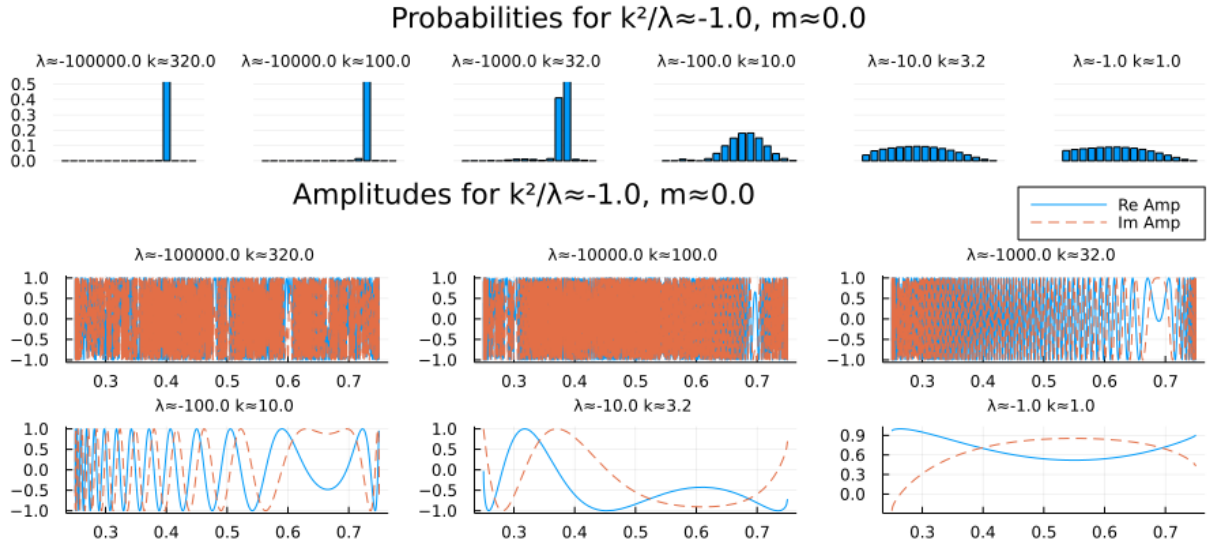

Figure 3.41: Probability and amplitude distributions for a family of (λ, k) with $k^2/\lambda = -1, m = 0$.

3.7 Light amplitudes and the continuum limit

The previous results are approximations based on a coarse simplicial lattice. To improve the approximation the lattice needs to be refined as the exact result is defined in the lattice refinement limit of (3.4). However, our current understanding on how to take the continuum limit in Lorentzian quantum gravity is incomplete (see [95] and [96] for related discussions on causal dynamical triangulation and spin-foams, respectively).

Therefore for further studies of light ray fluctuations we are faced with two tasks. The first is to understand light ray propagation in quantum spacetime with improved quantitative accuracy. The second is to understand the continuum limit of non-perturbative Lorentzian quantum gravity. It would be nice if the two tasks could be tackled together. In this section we discuss some thoughts along this line. The main idea is to treat light ray probabilities as the physical quantity to compare across different lattices in performing renormalization group type analysis.

3.7.1 Light amplitudes

Consider a bounded region of quantum spacetime crossed by a test light ray. Denote by b the gravitational boundary condition for the region, and by r the locations and directions of the light ray when it crosses the boundary. Note that a test light ray usually crosses the boundary more than once, so r contains a list of variables. The **light amplitude**

$$A[b,r] = \int_{b,r} \mathcal{D}g \; A[g] \tag{3.46}$$

is defined as the path integral over all gravitational configurations g compatible with the boundary conditions r and b.

From the light amplitudes one could derive light ray probabilities by taking the modulus square of the amplitudes, possibly after some coarse-grainings of the light ray variables r. This procedure is exemplified on a simple box lattice Γ in previous sections, where we computed the coarse-grained amplitudes

$$A_\Gamma[b, r_{in}, i] = \int_{r_{out}\in I_i} dr \; A_\Gamma[b,r] \tag{3.47}$$

for different outgoing light location intervals I_i under fixed incoming light location r_{in}, took the modulus square, and obtained the light ray probabilities.

3.7.2 Renormalization group

The next task is to push the study to finer lattices. In lattice field theories this is usually tackled in a renormalization group analysis.

As discussed in Section 1.7 of Montvay and Münster [97], there are two commonly adopted alternatives in lattice refinement for quantum lattice field theories. One could either fix the bare coupling constants and consider how the physical quantities such as the renormalized mass change as the lattice is refined, or fix the physical quantities and consider how the bare coupling constants change as the lattice is refined. In either case, the goal is to identify fixed points

as special places in the space of the couplings where both the bare coupling constants and the physical quantities remain unchanged as the lattice is refined. In principle, the theory space comes with infinitely many coupling constants, but in practice, one often works on a subspace with a finite number of couplings. For theories exhibiting asymptotic safety, fixed points that capture the essential aspects of the full theory can be identified in the finite-dimensional subspace.

For example in scalar field theory, one works in a two-dimensional theory space of the bare mass and the bare quartic couplings [97]. The renormalized mass and the renormalized quartic coupling relate more directly to laboratory observations and are picked as the physical quantities to compare across different lattices. Lattice refinement is carried out by decreasing the lattice spacing. In 4D one identifies the free theory fixed point but no interacting fixed points, which could be read as an indication that before reaching the zero lattice spacing limit some new ingredients such as additional degrees of freedom or a fundamental cutoff must come in to rescue the existence of an interacting quantum scalar field theory [98].

For simplicial quantum gravity one could attempt a similar study. Start with an ansatz for a finite-dimensional theory space. Refine the lattice by enlarging the lattice graph, which at the expectation value level decreases the simplicial volumes $\langle |V_s| \rangle$ when of the total spacetime volume $\langle \sum_s |V_s| \rangle$ is bounded.

The important question is what to pick as the physical quantities to compare across lattices. Lorentzian quantum gravity reveals light ray probability as a candidate. In previous sections we computed the coarse-grained light amplitudes and probabilities

$$A_{\Gamma,\alpha}[b, r \in I_i], \quad p_{\Gamma,\alpha}[b, r \in I_i] \tag{3.48}$$

for different bare couplings α on a simple lattice Γ under symmetry-reduction. To proceed, one should try to identify fixed points in the α theory space where $p_{\Gamma,\alpha}$ remains unchanged as the lattice Γ is refined. If successful, the constant physics trajectories approaching the fixed points in the theory space indicate how the continuum limit can be taken.

In this program it is crucial to find an efficient method of computation. In addition to numerical methods such as the one in [2], analytical insights including ideas on how to simplify the theory without sacrificing key aspects could be helpful. Note also that the light amplitude is composable in the sense that the light amplitude on the union of multiple quantum spacetime regions can be derived from summing the products of light amplitudes for the individual regions. Combined with the scaling identities of Section 3.2.4, it may be possible to find some shortcuts in computing light amplitudes on refined lattices by iteratively composing light ray amplitudes on coarser lattices.

3.7.3 Elementary light ray fluctuations?

In considering the strict infinite lattice limit a question arises on light ray fluctuations in the elementary simplicies. Should we regard an elementary simplex as a flat region of spacetime where light ray fluctuations are absent, or should we regard it as representing a family of spacetime configurations in superposition where light ray fluctuations are present?

In order to satisfy quantum uncertainty relations path integrals are dominated by non-differentiable configurations (see Section 7.3 of Feynman and Hibbs [13]). For example, take the path integral

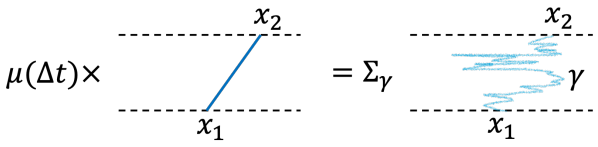

Figure 3.42: The path integral measure factor μ relates one linear path segment amplitude to the sum of amplitudes over a family of paths.

of a non-relativistic quantum particle and suppose to the contrary that the paths are all differentiable. Then for a fixed path at a time t the position $x(t)$ and momentum $p(t)$ are well-defined real numbers so that $x(t)p(t) - p(t)x(t) = 0$ for this path. Since this holds for all paths,

$$\langle x(t)p(t) - p(t)x(t)\rangle = 0 \tag{3.49}$$

as computed from the path integral. Quantum uncertainty relations would be violated, and one concludes that the assumption that the paths are differentiable cannot hold. On the other hand, we standardly enumerate the paths in terms of piecewise linear paths [13]. Along the interior of each linear piece of path, both the position and the momentum are well-defined real numbers. Is there not a contradiction?

One resolution is to interpret a piecewise linear path as representing a family of paths including non-differentiable ones (Figure 3.42). In the in the continuum (infinitely many time steps) limit the path integral measure factor serves to correct the difference between the amplitude of a piecewise linear path and the amplitude of family of paths including non-differentiable ones. Namely as the time interval between two steps $\Delta t \to 0$,

$$\mu(\Delta t)A[\gamma_0] = \sum_\gamma A[\gamma] \tag{3.50}$$

as illustrated in Figure 3.42 for one linear piece of path. Here $A[\gamma_0]$ is the amplitude for the linear path γ_0 connecting the starting and ending points, μ is the measure factor, and the RHS sums over paths γ including non-differentiable ones.

For example for a free particle with mass m, the transition amplitude from X_1 to X_2 in time T is

$$\sqrt{\frac{m}{2\pi i T}}e^{im(X_2-X_1)^2/2T}, \tag{3.51}$$

which can be obtained for instance from solving the Schrödinger equation. Had we approximated the transition amplitude by the amplitude of the linear path from X_1 to X_2, it would be $e^{im(X_2-X_1)^2/2T}$, which equals the second factor of (3.51). This approximation misses all the other paths connecting X_1 to X_2, including the non-differentiable ones. Therefore the first factor of (3.51) can be understood as correcting the differences in the time interval T. Now if we evaluate the same transition amplitude over time T using the standard path integral prescription over piecewise linear paths, during a time interval Δt a linear piece of path has amplitude $A[\gamma_0] = e^{im(x_2-x_1)^2/2\Delta t}$ between some positions x_1 and x_2. This would eventually lead to the wrong result in the $\Delta t \to 0$ limit because we missed all the other paths connecting x_1 to x_2, including the non-differentiable ones. Since (3.51) applies to all time durations, correcting the difference requires the multiplication of $\sqrt{\frac{m}{2\pi i\Delta t}}$, equals the standard path integral measure factor

$\mu(\Delta t)$. Hence the correct result is obtained by interpreting a piecewise linear path as representing a superposition over a family of paths whose amplitude sum differs from amplitude of the piecewise linear path by the measure factors.

A similar story can be told about simplicial gravitational path integrals which generalize the one-dimensional piecewise linear paths to higher dimensional piecewise flat simplicies. In analogy with the particle case, an elementary simplex of a simplicial spacetime configuration multiplied by the measure factor can be thought of as representing the sum over a family of configurations dominated by non-differentiable ones.

Therefore in computing light amplitudes on a finite lattice, it is reasonable to assume the presence of "elementary light ray fluctuations". Namely, even on an elementary simplex the light ray locations should not be sharply peaked as on a piece of Minkowski spacetime, but should exhibit quantum fluctuations due to the sum over a family of configurations. This means the light amplitudes for the elementary simplices should be non-vanishing for a range of light ray locations.

On the other hand, it is not clear to what extent this point is practically relevant in studying the continuum limit. It could be that once the lattice is taken large enough so that the elementary simplicies are taken small enough (in the sense of having small expectation value for the volume), introducing elementary light ray fluctuations or not does not affect the results. Further studies are needed to see if this is the case.

3.8 Discussion

In this work we studied light ray fluctuations in a simple model in Lorentzian simplicial quantum gravity. The overall question is how a quantum region of spacetime affects the propagation of light rays crossing it. We computed the probabilities for test light rays to land at different locations across a symmetry-reduced box model with simple boundary conditions in 2,3 and 4 spacetime dimensions.

For fixed boundary conditions light ray fluctuations are generically large when all coupling constants are relatively small in absolute value in all dimensions. In contrast, for fixed coupling constants light ray fluctuations show different trends in different dimensions when the boundary size is decreased. While in 2D light ray fluctuations first increase and then decrease, in 3D and 4D light ray fluctuations just increase without decreasing. The difference can be understood by noting that the coupling constants of the cosmological constant, Einstein-Hilbert, and R^2-terms have different length dimensions in different spacetime dimensions, so behave differently as the boundary length is scaled as in Section 3.2.4.

The symmetry-reduced box models with simple boundary conditions can be generalized in two directions. Firstly we could allow more dynamical degrees of freedom by relaxing the symmetry-reduction assumption and/or introducing larger simplicial lattices. This generalization is a necessity if we are to investigate the continuum limit of the theory as discussed in Section 3.7 where we propose to explore the use light amplitudes in renormalization group type analysis for Lorentzian quantum gravity. The continuum limit in turn would allow us to fix the coupling constants by comparing with empirical data (see [22] for a related discussion in Euclidean simplicial quantum gravity) in order to turn the qualitative conclusions about the amount of light ray fluctuation into quantitative predictions. Secondly we could consider additional boundary conditions.

Of particular interest are tunneling boundary conditions that admit complex, non-Lorentzian stationary points which are of interest to singularity resolution (see [2] and references therein). Because the path integrals are not dominated by particular Lorentzian configurations, light ray fluctuations could be generically large for singularity resolving tunneling processes.

Chapter 4

Truly Lorentzian quantum cosmology

Quantum cosmology based on Lorentzian path integrals is a promising avenue. However, many previous works allow non-Lorentzian configurations by integrating the squared scale factor over the whole real line. Here we show that restricting the minisuperspace path integral to Lorentzian configurations with positive squared scale factor can significantly change the expectation values. In addition, this enables the study of causal horizons and their quantum fluctuations, and achieves singularity avoidance trivially by excluding singular minisuperspace geometries as non-Lorentzian. The results indicate that semiclassical saddle point approximation is not always valid in truly Lorentzian quantum cosmology. As a consequence, related works on the tunnelling and no-boundary proposals, bouncing cosmology, and the quantum origin of inflation etc. need to be reexamined.

4.1 Introduction

Transitioning to the Lorentzian signature has been a recurring theme in quantum gravity. Historically, approaches such as simplicial quantum gravity [18], dynamical triangulation [99], spinfoam and related gauge theories [100] started in the Euclidean. Subsequent works eventually encompassed the Lorentzian setting to counter issues such as spikes [23, 5], degenerate geometries [101, 102], conformal instabilities [103], or to simply engage with *spacetime* which is Lorentzian.

In quantum cosmology one studies simplified models of quantum gravity such as the homogeneous and isotropic minisuperspace model with the metric

$$\mathrm{d}s^2 = -N^2\mathrm{d}\tau^2 + a(\tau)^2\mathrm{d}\Omega^2, \tag{4.1}$$

where $\mathrm{d}\Omega^2$ is the metric of a closed spatial 3-sphere. Since Euclidean gravitational path integrals suffer from the conformal instability issue [20], old works explored different complex integration contours [72, 104, 56, 105, 106, 107, 108, 109, 110, 111]. There have been various discussions about fixing the integration contour to be over Lorentzian spacetimes in the past (e.g., [112, 113, 114, 115]). More recently, Feldbrugge *et al.* proposed [60] to define the gravitational path integrals by the Lorentzian contour, and use Picard-Lefschetz theory to study complex contour deformations only as a computational trick for the fundamentally Lorentzian theory (see Sorkin [116] for a closely related discussion).

This has led to renewed interest in investigating old topics with new methods [60, 61, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138]. In these works of "Lorentzian quantum cosmology ", it is common to adopt the minisuperspace metric

$$\mathrm{d}s^2 = -\frac{N^2}{q(t)}\mathrm{d}t^2 + q(t)\mathrm{d}\Omega^2, \tag{4.2}$$

and treat the squared scale factor q as a path integral variable. The action in q is quadratic and a Gaussian integration yields nice closed-form results [139, 105].

However, the Gaussian integration over all real values of q is questionable step. For the metric to stay in the Lorentzian signature, q should only assume positive values. A path integral over also negative q cannot be said to be truly Lorentzian. In the seminal paper [139] that adopted q as a path integration variable, Halliwell gave ample discussions about potential issues for integrating the squared scale factor over all real values. Yet these warnings are largely left aside in subsequent works.

In this work we consider quantum cosmology in a truly Lorentzian setting, where the path integral is only over strictly Lorentzian configurations. In particular, we study minisuperspace quantum cosmology based on the metric (4.2), and distinguish the **real q scheme**, where q is integrated over the whole real line, from the **positive q scheme**, where q is integrated over positive values.

We find that the two schemes differ in at least three important aspects. First, the expectation values for the squared scale factor can differ much in the two schemes, when q gets close to or below zero for a relevant saddle point of the path integral. This affects the studies of tunnelling and no-boundary proposals, bouncing cosmology, and the quantum origin of inflation. Second, it is only possible to study the causal horizons and their quantum fluctuations in the positive q scheme. In the real q scheme the path integral includes non-Lorentzian geometries, where causal horizon is not defined. Third, restricting the path integral to the Lorentzian implies singularity avoidance. This is because singular minisuperspace geometries are non-Lorentzian and hence are automatically excluded from the path integral. In this sense, singularity avoidance is trivially achieved [8] in the truly Lorentzian minisuperspace path integral.

The results challenge the universal validity of semiclassical saddle point approximation. In particular, for negative spatial curvature bouncing cosmology, our numerical results based on the generalized thimble method [33] show that the saddle point which dominates the real q scheme path integral completely fails to capture the quantum expectation values of the truly Lorentzian positive q scheme. The true expectation values rather resembles that of the zero spatial curvature case in their real parts, in addition to possessing a large imaginary part. Here neither real nor complex (tunnelling) solutions to Einstein's equations characterize the path integral at leading order, because neither does the saddle point belong to the Lorentzian integration contour, nor does it connect to any configuration of the Lorentzian contour through the Picard-Lefschetz holomorphic gradient flow. As a consequence, semiclassical saddle point approximation should only be applied when its validity can be ascertained.

The paper is organized as follows. In Section 4.2 we review recent works on Lorentzian quantum cosmology. In Section 4.3 we point out the limitations of the real q scheme. In Section 4.4 we review the generalized thimble method which we use for numerical computation. In Section 4.5 we define the quantities of lightcone location and its fluctuations to be computed. In Section 4.6

we put the pieces together to study bouncing cosmology for positive, zero, and negative spatial curvatures, and compare results from the positive and real q schemes. In Section 4.7 we discuss singularity avoidance. In Section 4.8 we conclude with a discussion of some topics for further study.

In the following, we set $c = \hbar = 8\pi G = 1$.

4.2 Previous works

Following [60, 139, 105] we consider the minisuperspace metric

$$\mathrm{d}s^2 = -\frac{N^2}{q(t)}\mathrm{d}t^2 + q(t)\left(\frac{1}{1-kr^2}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2\right)\right), \tag{4.3}$$

with squared scale factor $q(t)$, lapse N, and spatial curvature $k = 1, 0$ or -1.[1] The $dN/dt = 0$ gauge is used so N does not depend on t.

Plugging (4.3) in the Einstein-Hilbert action with the Gibbons-Hawking-York boundary term [140, 141] $S = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}(R - 2\Lambda) + \int_B \mathrm{d}^3y\sqrt{h}K$, one obtains

$$S[q, N] = 2\pi^2 \int_0^1 \mathrm{d}t\, N\left(-\frac{3\dot{q}^2}{4N^2} - \Lambda q + 3k\right). \tag{4.4}$$

The dot denotes derivative with respect to the coordinate time t, which is taken to run from 0 to 1. This is without loss of generality, since the physical proper time derived from (4.3) is still arbitrary due to N and q.

For the boundary condition $q(0) = q_0, q(1) = q_1$ the path integral takes the form

$$Z[q_0, q_1] = \int DN \int_{q_0,q_1} Dq\; e^{iS[q,N]}. \tag{4.5}$$

We omit the subscript q_0, q_1 below when no ambiguity arises.

The metric (4.3) is written in terms of the squared scale factor rather than the scale factor $a(t)$. This produces the action (4.4) which is quadratic in q. In many previous works, such as [60, 61, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138], the integration range of the squared scale factor is *taken to be over the whole real line*. Then the path integral in q with the quadratic action can be evaluated analytically, just like the path integral of a free quantum particle [13].

Explicitly, a Gaussian functional integration with respect to q yields

$$G[q_0, q_1; N] := \int Dq\; e^{iS[q,N]} = \sqrt{\frac{3\pi i}{2N}} e^{i2\pi^2 \int_0^1 \mathrm{d}t(-\frac{3}{4N}\dot{\bar{q}}^2 + N(3k - \Lambda\bar{q}))}, \tag{4.6}$$

$$\bar{q}(t) = \frac{\Lambda}{3}N^2t^2 + \left(-\frac{\Lambda}{3}N^2 + q_1 - q_0\right)t + q_0. \tag{4.7}$$

[1] For the flat ($k = 0$) and hyperbolic ($k = -1$) cases, we assume that the spatial geometry is compactified as in [107] so that the action (4.4) is finite.

Here $\bar{q}(t)$ obeys the boundary condition $q(0) = q_0, q(1) = q_1$ and solves

$$\ddot{q} = \frac{2\Lambda}{3} N^2, \tag{4.8}$$

which is the equation of motion obtained from $\delta S/\delta q = 0$.

To obtain the final result $Z[q_0, q_1] = \int DN\ G[q_0, q_1; N]$ from (4.6), one still needs to analyze the N integral. In previous works this one-dimensional integral is commonly studied through a saddle point approximation. The saddle point can be obtained directly by demanding stationary phase for $G[q_0, q_1; N]$. Since the phase $2\pi^2 \int_0^1 \mathrm{d}t(-\frac{3}{4N}\dot{\bar{q}}^2 + N(3k - \Lambda\bar{q}))$ equals $S[\bar{q}, N]$, we have $\partial_N S[\bar{q}, N] = 0$. Equivalently, we could look at the original path integral (4.5) and demand $\delta S[q, N]/\delta N = 0$ to obtain the equation of motion

$$\int_0^1 \mathrm{d}t \left(\frac{3}{4N^2}\dot{q}^2 + 3k - \Lambda q \right) = 0. \tag{4.9}$$

Equation (4.9) and (4.8) form the complete set of equations of motion for the variables q and N. In the joint solution, q is given by (4.7), and N is given by

$$\bar{N} = c_1 \frac{3}{\Lambda} \left((\frac{\Lambda}{3} q_0 - k)^{1/2} + c_2 (\frac{\Lambda}{3} q_1 - k)^{1/2} \right), \tag{4.10}$$

where $c_1, c_2 \in \{-1, 1\}$.

This offers four possible saddle points. However, not all of them will make contributions to the Lorentzian path integral, and Picard-Lefschetz theory can be employed to determine the actually relevant saddle points [60]. Previous works show that the relevancy of the saddle points depends on whether the N-integral is defined as $\int_0^\infty \mathrm{d}N$ or $\int_{-\infty}^\infty \mathrm{d}N$. There has been no consensus in the literature on which measure to use [60, 61, 117, 118, 119, 120].

4.3 Limitations of the real q scheme

4.3.1 Cases with limitations

In the procedure reviewed above, integrating $q(t)$ over the whole real line is crucial. It enables Gaussian integration to obtain the analytic result (4.6).

However, in the context of *Lorentzian* quantum cosmology there is an unsettling issue. For the metric (4.3) to stay in the Lorentzian signature $(-,+,+,+)$, it must be that $q(t) > 0$. Therefore q should only be integrated over positive values in a strictly Lorentzian path integral.

In practice, integrating q over the real line could still be employed as a useful trick if the result agrees well with integrating over positive q. For instance, if the saddle point $\bar{q}$ of (4.7) stays far above zero for the whole time $t \in [0, 1]$, then the integrals in both positive and real q schemes are dominated by paths which stay positive.

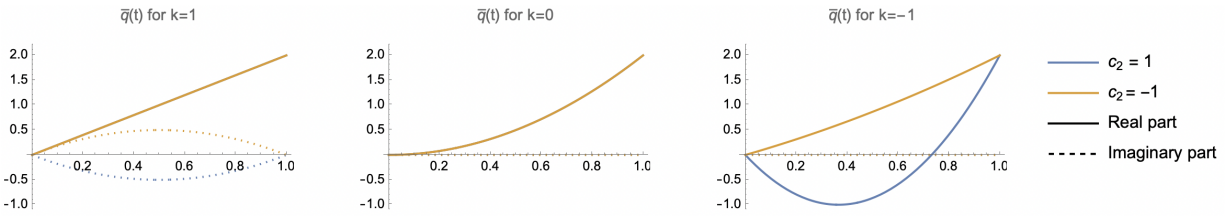

Figure 4.1: Plotting $\bar{q}(t)$ with $\Lambda = 3, q_0 = 0, q_1 = 2$ for on-shell N with $c_1 = 1$. In the first two cases $\mathrm{Re}\,\bar{q}$ overlap for $c_2 = 1$ and $c_2 = -1$.

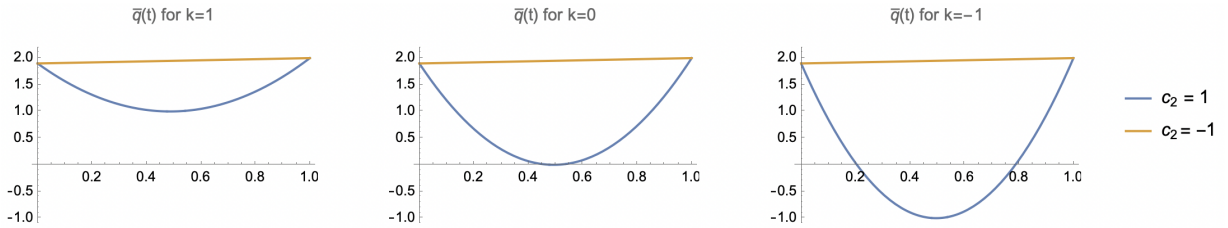

Figure 4.2: Plotting $\bar{q}(t)$ with $\Lambda = 3, q_0 = 1.9, q_1 = 2$ for on-shell N with $c_1 = 1$.

Yet in some cases $\bar{q}$ does not stay far above zero:

- When a boundary value q_0 or q_1 is close to zero, then (4.7) clearly does not stay far away from zero for all time (Figure 4.1). This happens for Lorentzian versions (e.g. [60, 61, 121, 125, 126]) of the tunnelling [43, 142, 143] and no-boundary [44] boundary conditions where q_0 is sent to zero.
- When N is on-shell at (4.10), the saddle point $\bar{q}$ can reach $3k/\Lambda$ (e.g., at the minimum value for the $c_2 = 1$ cases in Figure 4.2). This minimum value can get close to or below zero. For example, this happens in the $k = 0$ case relevant to inflation [123].
- When N is allowed off-shell, there are more cases where $\bar{q}(t)$ gets close to or below zero. For example, when $\Lambda > 0$ the bouncing saddle point $\bar{q}$ always dives into negative values for sufficiently large N.[2]

The first two cases are especially troublesome. Here the relevant saddle points with both q and N set on-shell get close to or below zero. Paths at and around these saddle points make significant contributions to the path integral in the real q scheme, but are excluded in the positive q scheme. Therefore the real q scheme result may deviate much from the truly Lorentzian positive q scheme result.

[2]The general solution (4.7) is a parabola with axis of symmetry at $t_a = 1/2 + 3(q_0 - q_1)/(2N^2\Lambda)$. When $\Lambda > 0$, $\bar{q}$ assumes its minimum value

$$\bar{q}(t_a) = \frac{1}{12}\left(-\Lambda N^2 - \frac{9(q_0 - q_1)^2}{\Lambda N^2} + 6(q_0 + q_1)\right), \tag{4.11}$$

which is always negative for large enough N. Moreover, t_a approaches $1/2$ for large N, so it always fall within the relevant range $t \in [0, 1]$.

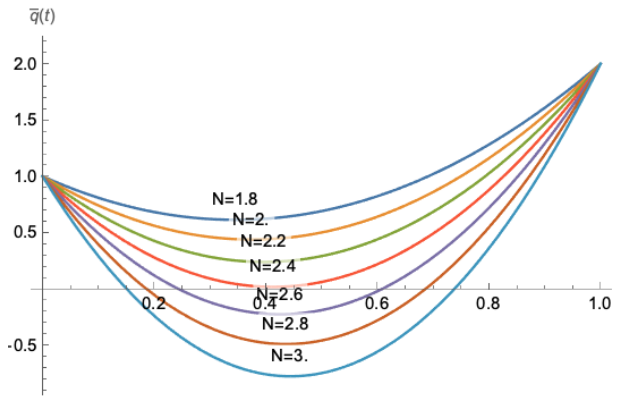

Figure 4.3: Plotting $\bar{q}(t)$ with $\Lambda = 3, k = 1, q_0 = 1, q_1 = 2$ for a list of N values.

- In addition, for any values of q_0, q_1, Λ, k, path integral configurations with zero or negative q at some time are not Lorentzian so do not possess a causal structure.

This poses a difficulty in studying topics related to causal structures, for example, on the topics of how quantum fluctuations of spacetime affects the horizon problem [3] and light/gravitational wave propagations for bouncing cosmology [144].

4.3.2 A toy model example

To illustrate how the positive and real q schemes can produce quantitatively very different results, we look at a simple toy model just for the q path integral. The results of the later sections, which demand more efforts to obtain, will show that the same happens for the joint $q-N$ path integral. The q path integral can be approximated by

$$G[q_0, q_1; N] \approx \int Dq\, e^{i\sum_{i=0}^{n} S_i}, \tag{4.12}$$

$$\int Dq = \int_0^\infty dq(t_1) \cdots \int_0^\infty dq(t_n)\, \mu\left(q(t_1), \cdots, q(t_n), N\right), \tag{4.13}$$

$$S_i = 2\pi^2 \left(-\frac{3\left(q(t_{i+1}) - q(t_i)\right)^2}{4N\Delta t} + N\Delta t \left(3k - \frac{1}{2}\Lambda \left(q(t_i) + q(t_{i+1}) \right) \right) \right), \tag{4.14}$$

where the time domain $t \in [0, 1]$ is broken into $n+1$ intervals of size $\Delta t = 1/(n+1)$ with the actions S_i. The exact result is approached as $n \to \infty$.

Here μ is the measure factor for the integrals. The result (4.6) is obtained with

$$\mu = \left(\frac{3\pi i}{2N\Delta t} \right)^{\frac{n+1}{2}}, \tag{4.15}$$

which takes the same form of the measure factor for a quantum particle [13]. We will use this measure to make the comparison between the positive and real q schemes.

Figure 4.3 plots the on-shell $\bar{q}(t)$ of (4.7) for $\Lambda = 3, k = 1, q_0 = 1, q_1 = 2$ for a list of N values, including some for which $\bar{q}(t)$ gets close to zero or reach negative values. In the simplest

	N=1.8	N=2.	N=2.2	N=2.4	N=2.6	N=2.8	N=3.
v1 for q>0	0.19 - 1.61 i	0.15 + 1.53 i	1.42 - 0.37 i	1.39 - 0.21 i	-0.57 + 1.22 i	1.29 + 0.16 i	0.98 + 0.78 i
v2 for q∈R	0.13 - 1.7 i	0. + 1.58 i	1.67 - 0.37 i	0.89 + 0.03 i	0.26 + 0.22 i	0.05 + 0.16 i	0.02 + 0.1 i
|v1-v2|/|v1|	0.07	0.1	0.18	0.39	0.96	0.95	0.94

Figure 4.4: Numerical integration results for the $n = 1$ approximation. The first row shows results of the positive q scheme, the second row shows results of the real q scheme, and the third row shows their relative differences.

approximation $n = 1$, there is only one dynamical variable $q := q(t_1)$, and (4.12) can be computed by direct numerical integration. The results obtained using *Mathematica* for the positive and real q schemes are shown in Figure 4.4. Clearly as N increases and $\bar{q}(t)$ approaches zero or negative values, the difference becomes quite significant. Already at $N = 2.4$ where $\bar{q}(t)$ still stays positive, the difference reaches as high as 39%.

4.4 Generalized thimble method

In order to investigate the differences between the truly Lorentzian positive q scheme and the real q scheme further, we need a method to evaluate the truly Lorentzian path integrals. The problem is quite non-trivial because analytically, not much is known for path integral computations beyond Gaussian integration. Even numerically, the complex Lorentzian path integral has an oscillating phase that gives rise to the numerical sign problem.

4.4.1 Review of the method

Fortunately, the generalized thimble numerical method [33, 34] offers a way to overcome the sign problem. This is a Monte Carlo sampling method that exploits Picard-Lefschetz theory to deform the integration contour to reduce the complex phase fluctuations. It can be viewed as a generalization of the Lefschetz thimble method [30] to other than the steepest descent contours [31], which makes the method more adaptable to attack problems such as multimodal problems [81, 80, 85, 145].

Given a multidimensional integral

$$\int \prod_i dv_i \; e^{E[v_1, v_2, \cdots]}, \tag{4.16}$$

the **holomorphic gradient flow** equations

$$\frac{dv_i}{dt} = -\overline{\frac{\partial E}{\partial v_i}} \quad \forall i \tag{4.17}$$

generates an integral curve for each point $\zeta = (v_1, v_2, \cdots, v_n)$ in the original integration contour. Subjecting the whole integration contour to this flow generates a contour deformation $C(t)$ as a function of the flow time t, with $C(0)$ as the original contour. If the integrand is holomorphic everywhere the flow transverses, Cauchy's integration theorem guarantees that the integral along $C(t)$ differs from the original one only along the boundaries of the flowed region (Figure 4.5). If

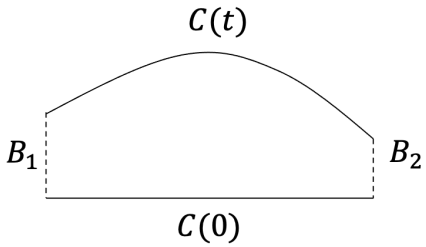

Figure 4.5: If the integrand is holomorphic in the region enclosed by the curves shown, the integral along the boundary will vanish by Cauchy's integration theorem. As a consequence, the integrals along the contours $C(0)$ and $C(t)$ differ only by the integrals along the dashed boundaries.

the boundary contributions are negligible, we could use the integral along $C(t)$ to approximate the original integral.

Evaluating the integral along $C(t)$ could ameliorate the sign problem. To see this, note that by (4.17) the real part E_R of E obeys

$$\frac{\mathrm{d}E_R}{\mathrm{d}t} = \frac{1}{2}\left(\frac{\mathrm{d}E}{\mathrm{d}t} + \overline{\frac{\mathrm{d}E}{\mathrm{d}t}}\right) = \frac{1}{2}\sum_i \left(\frac{\partial E}{\partial v_i}\frac{\mathrm{d}v_i}{\mathrm{d}t} + \overline{\frac{\partial E}{\partial v_i}\frac{\mathrm{d}v_i}{\mathrm{d}t}}\right) = -\sum_i \left|\frac{\partial E}{\partial v_i}\right|^2 \leq 0. \tag{4.18}$$

Therefore the magnitude of the integrand is exponentially suppressed along the flow, except for regions close to the stationary points where $\partial E/\partial v_i = 0, \forall i$. For sufficiently large t, only this region contributes significantly to the integral along $C(t)$, and we can hope that the phase fluctuation of the integrand is milder than over the original contour.

The generalized thimble method of [33, 34] exploits the deformed contour to perform Markov Chain Monte Carlo sampling based on the following algorithm:

1. Fix some flow time $t = T$. Start with a configuration $\zeta = \{v_i\}_i$ in the original contour. Use numerical integration to evolve it under (4.17) by T to obtain $\phi = \phi(\zeta)$.

2. Sample a new configuration $\zeta' = \zeta + \delta\zeta$ on the original contour and evolve ζ' under (4.17) again by T to obtain $\phi' = \phi'(\zeta')$.

3. Accept ζ' as the new ζ with probability $P = \min\{1, e^{\mathrm{Re}\, E_{\mathrm{eff}}(\phi') - \mathrm{Re}\, E_{\mathrm{eff}}(\phi)}\}$, where E_{eff} is defined below in (4.22).

4. Repeat steps 2 and 3 until a sufficient ensemble of configurations is generated.

5. Compute the expectation values using the formula

$$\langle O \rangle = \frac{\left\langle O e^{i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}}{\left\langle e^{i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}}, \tag{4.19}$$

   where φ is defined in (4.23) and $\langle\cdot\rangle_{\mathrm{Re}\, E_{\mathrm{eff}}}$ denotes averaging over the ensemble just generated.

To define E_{eff}, we note that

$$\int_{C(0)} e^{E(\zeta)} d\zeta = \int_{C(T)} e^{E(\phi)} d\phi = \int_{C(0)} e^{E(\phi(\zeta))} \det J \, d\zeta. \tag{4.20}$$

In the second expression, the contour $C(T)$ is parametrized by the flowed coordinates ϕ. In the last expression the contour $C(T)$ is reparametrized by the original coordinates ζ. This induces the Jacobian $J_{ij} = \frac{\partial \phi_i}{\partial \zeta_j}$ which can be computed by integrating to $t = T$

$$\frac{dJ_{ij}(t)}{dt} = \sum_k \overline{H_{ik} J_{kj}}, \quad H_{ij} = -\frac{\partial^2 E}{\partial v_i \partial v_j}, \quad J_{ij}(0) = \delta_{ij}. \tag{4.21}$$

The integrand exponent of the last integral of (4.20) is given a special name

$$E_{\text{eff}} = E(\phi(\zeta)) + \log \det J(\zeta) \tag{4.22}$$

Expanding E_{eff} in real and imaginary parts yields $e^{E_{\text{eff}}} = e^{\operatorname{Re} E_{\text{eff}} + i\varphi}$, where

$$\varphi = \operatorname{Im} E_{\text{eff}} = \operatorname{Im} E + \arg \det(J). \tag{4.23}$$

This explains steps 3 and 5, in which we sample (4.20) according to the magnitude $e^{\operatorname{Re} E_{\text{eff}}}$ of the integrand, and multiply O with the phase $e^{i\varphi}$ in (4.19).

4.4.2 Integration range and measure factors

We want to apply the generalized thimble method to the path integral (4.5)

$$Z[q_0, q_1] = \int DN \int Dq \, e^{iS[q,N]} = \int DN \, G[q_0, q_1; N] \tag{4.24}$$

with $G[q_0, q_1; N]$ given in (4.12). For this we need to specify the integration range and measure factors.

Since the metric (4.3) is of the Lorentzian signature $(-,+,+,+)$ only when q is positive, we integrate q over positive values as in (4.13).

For the q measure factor μ of (4.13), previous results in the real q scheme employed (4.15). Since $\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\pi/a} = 2\int_0^{\infty} e^{-ax^2} dx$ for $\operatorname{Re} a > 0$, it seems reasonable to modify (4.15) by a constant factor in the positive q scheme. Since constant multiplicative factors cancel out in (4.19), for simplicity we will directly employ $\mu = (\frac{1}{N})^{\frac{n+1}{2}}$. This factor μ can be incorporated as an additional term

$$E_\mu = \frac{n+1}{2} \log N \tag{4.25}$$

in the path integral exponent E of (4.16).

As mentioned at the end of Section 4.2, there is more than one choice for the N integration range. Here we take

$$\int DN = \int_0^{\infty} \mathrm{d}N \tag{4.26}$$

as in [60] for concreteness. The disagreement between real and positive q schemes should also be present for the alternative measure $\int_{-\infty}^{\infty} \mathrm{d}N$. Although we have not performed the study, it seems the generalized thimble method can be applied to this case as well.

An additional measure factor is included for the following reason. The q and N integration ranges are both bounded by 0. As explained around Figure 4.5, the integrals along the flowed contour and the original contour agree well if contributions are small along the "side contours" (dashed part of Figure 4.5) traced by the boundaries of the original contour under the holomorphic gradient flow. For the particular case at hand, the boundaries of the original lie at $N = 0$ and $q(t_i) = 0,\ \forall i$. According to (4.25), the $N = 0$ boundary is a log singularity. Since the exponent E is not holomorphic there, the $N = 0$ boundary of the original contour lies outside the domain of the holomorphic gradient flow, and is unchanged under the flow. Therefore this "side contour" has no extension, and offers no contribution to the integral.

To ensure the same at the $q(t_i) = 0$ boundaries, we employ the trick of introducing the measure factor $\prod_i q(t_i)^m$ for some $m < 0$. For $|m| \ll 1$ such as $m = -0.001$ used here, q^m stays fairly close to one for practical ranges of q. For instance from $q = 10^{-8}$ to $q = 10^2$, q^m decreases monotonically from 1.0186 to 0.9954 to four digits after the decimal place. Yet as in the N case, it generates a log singularity for the exponent, which puts the $q(t_i) = 0$ boundaries outside the domain of the holomorphic gradient flow. Consequently, the $q(t_i) = 0$ boundaries are unchanged under the flow, such that the "side contour" has no extension, and offers no contribution to the integral. This ensures that the integrals agree along the flowed contour and the original contour.

In summary, we will apply the generalized thimble method to the path integral

$$Z[q_0, q_1] = \int_0^\infty \mathrm{d}N \int_0^\infty dq(t_1) \cdots \int_0^\infty dq(t_n)\ e^{i \sum_{i=0}^n \tilde{S}_i - \frac{n+1}{2} \log N}, \tag{4.27}$$

$$\begin{aligned} \tilde{S}_i =& 2\pi^2 \left(-\frac{3(q(t_{i+1}) - q(t_i))^2}{4N\Delta t} + N \left(3k - \frac{1}{2}\Lambda(q(t_i) + q(t_{i+1})) \right) \Delta t \right) \\ & - \frac{im}{2} \left(\log q(t_i) + \log q(t_{i+1}) \right), \end{aligned} \tag{4.28}$$

where the measure factor for q is absorbed in $\tilde{S}_i$, and that for N is added to the exponent of the integrand. For convenience of writing, we separated a single factor $q(t_i)^m$ into two places in $\tilde{S}_i$ and $\tilde{S}_{i-1}$. This introduces constant factors for the unintegrated boundary q values, but these constants drop out eventually when taking ratios as in (4.19). The integrals in (4.27) are now for the Borel measure without additional factors, so (4.17) is directly applicable.

4.4.3 Notes on implementing the algorithm

The generalized thimble method requires the integrand to be holomorphic everywhere the holomorphic gradient flow transverses. Since log functions show up in the integrand (4.27), the integration domain is now taken on the Riemann surfaces of the q's and N. This means in step 1 of the generalized thimble algorithm, one needs to keep track of the log branches for q and N during the flow. In the Julia programming language [84] that we use, this is implemented by the "callback functions" of the package "DifferentialEquations.jl" [146], as is done in simplicial quantum gravity which refers to both log and square root branches [2].

For step 2 of the generalized thimble algorithm, again as in [2] we apply the adaptive Monte Carlo sampler reviewed in [83]. The rest of the steps are then implemented as stated in Section 4.4.1.

4.5 Lightcone fluctuations

The results of the generalized thimble method are in terms of the expectation values (4.19). We are interested in $\langle q(t)\rangle$ and $\langle N\rangle$ for the squared scale factor and the lapse.

In addition, we will compute the expectation values for the lightcone location and their fluctuations. For the metric (4.3), the equation for radial lightlike geodesic is

$$0 = -\frac{N^2}{q(t)}\mathrm{d}t^2 + \frac{q(t)}{1-kr^2}\mathrm{d}r^2. \tag{4.29}$$

With $\mathrm{d}\chi^2 = \frac{1}{1-kr^2}\mathrm{d}r^2$, (4.29) implies $\frac{N\mathrm{d}t}{q(t)} = \mathrm{d}\chi$. Integrating both sides yields

$$N\int \frac{1}{q(t)}\mathrm{d}t = \Delta\chi. \tag{4.30}$$

During a time interval Δt, the zigzagging path of (4.12) obeys $q(t) = \frac{q(t_{i+1})-q(t_i)}{\Delta t}t + q(t_i)$. Plugging this in (4.30) for yields

$$\Delta\chi_i = \frac{N\Delta t(\log q(t_i) - \log q(t_{i+1}))}{q(t_i)-q(t_{i+1})}, \quad \Delta\chi = \sum_{i=0}^{n}\Delta\chi_i. \tag{4.31}$$

For a radial geodesic $\mathrm{d}s^2 = -\frac{N^2}{q(t)}\mathrm{d}t^2 + q(t)\mathrm{d}\chi^2$, so $\Delta\chi$ gives the spatial comoving distance that a radial light ray covers from $t=0$ to $t=1$ and quantifies the size of the causal horizon for events at $t=0$. Below we will use (4.31) in (4.19) to compute the expectation value $\langle\Delta\chi\rangle$ and the standard deviation

$$\sigma = \sqrt{\langle\Delta\chi^2\rangle - \langle\Delta\chi\rangle^2} \tag{4.32}$$

to quantify horizon fluctuations.

4.6 Case study: bouncing cosmology

In Figure 4.2 the bouncing saddle points $\bar{q}(t)$ are for $c_2 = 1$. When $k=1$, $\bar{q}(t)$ stays above 0. When $k=0$, $\bar{q}(t)$ reaches 0 at its minimum. When $k=-1$, $\bar{q}(t)$ drops below 0. In the last two cases, the saddle point $\bar{q}(t)$ does not stay positive for all time, so $\bar{q}(t)$ is not included in the truly Lorentzian path integral sum. Therefore results from the real q scheme run the risk of deviating much from the positive q scheme. In this section we apply the method of Section 4.4 to make a quantitative comparison between the positive and real q schemes.

4.6.1 Focusing on the bouncing saddle point

As shown in [60], both saddle points with $c_2 = 1$ and $c_2 = -1$ in Figure 4.2 are relevant for the real q scheme path integrals. An otherwise unconstrained path integral will exhibit interference effects for the two saddle points.

In comparing the positive and real q schemes, we want to focus on the $c_2 = 1$ bouncing saddle point since the $c_2 = -1$ saddle point stay high above 0 for all time. One way to achieve this is to

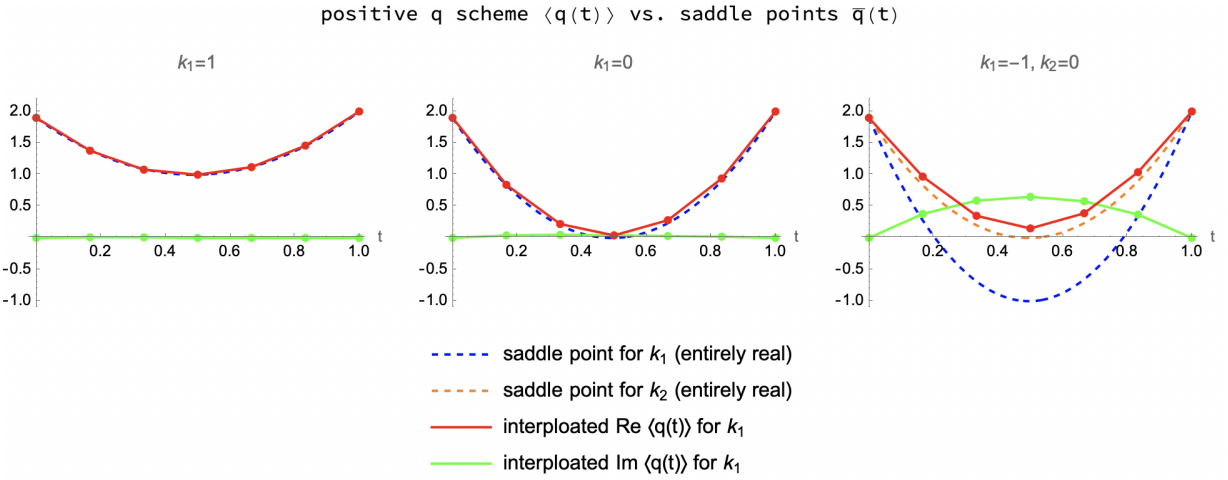

Figure 4.6: Comparing quantum expectation values $\langle q(t)\rangle$ of Table 4.1 from the positive q scheme with the saddle point values $\bar{q}(t)$.

modify the boundary condition. The $c_2 = 1$ and $c_2 = -1$ saddle points have different momentum at the boundaries. By employing a coherent state type boundary condition that centers around the momentum of $c_2 = 1$, one obtains a path integral without the interference of the other saddle point [123].

The generalized thimble method offers a practical alternative. In a Monte Carlo simulation with multiple saddle points, the sampler needs to overcome the low integrand weight barrier at intermediate regions to move from around one saddle point to around another. Usually one wants to sample efficiently across different saddle points, and some advanced variations of the original generalized thimble method have been developed to achieve this [80, 81, 85, 145]. Here we want to avoid travelling across saddle points, which does not require any advanced method. Because of (4.18), a larger flow time T increases the integrand weight barrier. Therefore we can simply adopt a large T to restrict the Monte Carlo sampler to around the bouncing saddle point. The results presented next show the appropriate T to achieve this.

4.6.2 Results

In Figure 4.2 we took $q_0 = 1.9, q_1 = 2.0$ and $\Lambda = 3$. For the same parameters, Figure 4.6 and Table 4.1 summarize the numerical results based on (4.27) with $n = 5$.

For $k = 1$ and $k = 0$, the expectation values $\langle q(t)\rangle$ and $\langle N\rangle$ from the positive and real q schemes are close, as seen in Table 4.1. The expectation values $\langle q(t)\rangle$ are in turn close to the saddle point values $\bar{q}(t)$, as seen from Figure 4.6. In particular, the imaginary part of the saddle points vanish, and the imaginary part of the expectation values are also close to zero. In contrast:

- For $k = -1$, the expectation values $\langle q(t)\rangle$ and $\langle N\rangle$ from the positive and real q schemes differ much. In the positive q scheme, $\mathrm{Re}\,\langle q(t)\rangle$ deviate much from $\bar{q}$ which is real, and $\mathrm{Im}\,\langle q(t)\rangle$ deviate much from zero.

Besides, there is a fundamental difference that applies to all values of k. In the positive q scheme all path integral configurations are Lorentzian, so we can compute expectation values

k	1	1	0	0	-1	-1
scheme	positive	real	positive	real	positive	real
m	-0.001	0	-0.001	0	-0.001	0
T	0.025	0.025	0.03	0.03	0.027	0.03
$\langle e^{i\varphi} \rangle$	1	1	0.98	0.98	0.94	0.96
$\langle q(t_1) \rangle$	1.38 + 0.01im	1.41 - 0.01im	0.84 + 0.04im	0.86 + 0.02im	0.97 + 0.38im	0.22 - 0.01im
$\langle q(t_2) \rangle$	1.08 + 0.01im	1.1 + 0.0im	0.22 + 0.05im	0.24 + 0.03im	0.35 + 0.59im	-0.77 - 0.02im
$\langle q(t_3) \rangle$	1.0 + 0.0im	1.0 + 0.02im	0.04 + 0.04im	0.03 + 0.04im	0.15 + 0.65im	-1.1 - 0.02im
$\langle q(t_4) \rangle$	1.12 - 0.0im	1.1 + 0.03im	0.28 + 0.03im	0.25 + 0.05im	0.39 + 0.58im	-0.74 - 0.01im
$\langle q(t_5) \rangle$	1.46 - 0.0im	1.44 + 0.03im	0.94 + 0.02im	0.91 + 0.04im	1.04 + 0.37im	0.29 - 0.01im
$\langle N \rangle$	1.96 - 0.03im	1.92 - 0.03im	2.7 - 0.11im	2.71 - 0.1im	1.73 - 0.81im	3.5 + 0.03im
$\langle \Delta\chi \rangle$	18.93 - 0.47im	N/A	97.96 - 37.04im	N/A	5.6 - 28.43im	N/A
σ	1.08 - 0.08im	N/A	11.84 + 18.21im	N/A	1.07 - 1.9im	N/A
$\lvert\sigma\rvert$	1.08	N/A	21.72	N/A	2.18	N/A

Table 4.1: Results for $q_0 = 1.9, q_1 = 2.0, \Lambda = 3$. Data for each of the six columns is produced from a Monte Carlo chain of length 10 million.

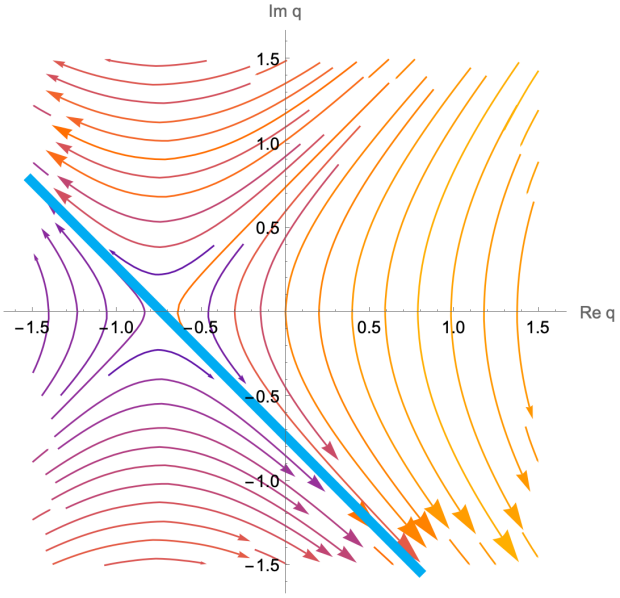

Figure 4.7: Holomorphic gradient flow for the one-variable model for $N = 3$. The steepest descent contour is labelled by the thickened line.

$\langle \Delta\chi \rangle$ for the comoving horizon (4.31) and its fluctuations σ as defined in Section 4.5. In contrast, in the real q scheme there are path integral configurations which are not Lorentzian, so the expectation value for the comoving horizon is undefined. The results for $\langle \Delta\chi \rangle, \sigma, |\sigma|$ for the real q scheme are shown in the last three columns of Table 4.1. To the extent that $|\sigma|$ offers an indirect indicator of the amount of fluctuation:

- In the positive q scheme, the amount of lightcone fluctuation as indicated by $|\sigma|$ is much larger for $k = 0$ than for $k = 1$ and $k = -1$.

4.6.3 Understanding the expectation values

In the first bullet point of Section 4.6.2, we noted that the real q scheme expectation values for q differ much from the saddle point values in the negative spatial curvature $k = -1$ case. To un-

derstand this, it helps to consider again the one-variable toy model of Section 4.3.2. According to Picard-Lefschetz theory [60], the real line approaches the steepest descent contour asymptotically under the holomorphic gradient flow defined by (4.16) and plotted in Figure 4.7.

In the real q scheme, the integration contour is the real line. Since this contour is deformed into the steepest descent contour under the flow, and the integral can be equivalently performed there. As shown in Figure 4.7, the saddle point is at -0.75. Points around the saddle point along the steepest descent contour all have negative real parts around -0.75, but some have positive and some have negative imaginary parts. Therefore we expect $\langle q(t)\rangle$ to have a negative real part around the saddle point, and an almost vanishing or exactly vanishing imaginary part due to the cancellation from positive and negative contributions.

In the positive q scheme, the integration contour is the positive halfline. As shown in Figure 4.7, this contour only approaches a portion of the steepest descent contour quite far from the saddle point. In particular, all points of the original contour flow towards the directions of positive real values and negative imaginary values. Therefore we expect $\langle q(t)\rangle$ to have a much larger real part than the saddle point, and a much smaller imaginary part than zero.

The actual model for $k = -1$ we considered has more dynamical variables than one. However, similarly the real parts of $\langle q(t)\rangle$ deviate much from the negative saddle point values, and the imaginary parts of $\langle q(t)\rangle$ deviate much from zero (Figure 4.8). Presumably this is for the same reason that under the holomorphic gradient flow, the original positive q contour approaches only a portion of the steepest descent contour which does not cover the saddle point.

4.6.4 Understanding the fluctuations

Consider three Gaussian distributions with standard deviation 1, but centered around different locations $\mu = -2, 0$ and 2. Assume that only the portion along the positive real halfline $\mathbb{R}^+$ is relevant, and rescale the distributions so that they are normalize on $\mathbb{R}^+$. As shown in Figure 4.10, the smaller μ is, the more sharply the distribution is peaked. As a consequence, the standard deviation σ computed on $\mathbb{R}^+$ is smaller for smaller μ (Figure 4.10).

In the case of quantum cosmology, the real vs. positive q schemes of $k = -1$ is like $\mu = 2$ vs. $\mu = -2$. In the real q scheme, the original contour covers the whole steepest descent contour under the holomorphic gradient flow, and the magnitude of the integrand varies slowly around the saddle point to yield relatively large fluctuations. In the positive q scheme, the original contour only covers a portion of the steepest descent contour under the holomorphic gradient flow. This portion does not contain the saddle point, and the magnitude of the integrand varies much faster around the peak that is covered. Therefore the fluctuations in $\langle q(t)\rangle$ and $\langle N\rangle$ are smaller in comparison to the real q scheme (Figure 4.8 vs. Figure 4.9).

The positive q scheme of $k = -1$ vs. $k = 0$ is like $\mu = -2$ vs. $\mu = 0$. For $k = 0$, the saddle point is on the boundary of the original contour, just like the $\mu = 0$ case. Here we expect more fluctuation than $k = -1$ which is analogous to the $\mu = -2$ case. This explains why there is more fluctuation for $k = 0$ than for $k = -1$ measured by $|\sigma|$, as noted in the second bullet point of Section 4.6.2.

Finally, there is also less fluctuation for $k = 1$ than for $k = 0$ as measured by $|\sigma|$. This is presumably because of the form $\Delta\chi_i \propto \frac{\log q - \log q'}{q - q'}$ of the comoving distance function (4.31). As q approaches zero, $\Delta\chi_i$ becomes very large and very sensitive to the precise value of q. Since in

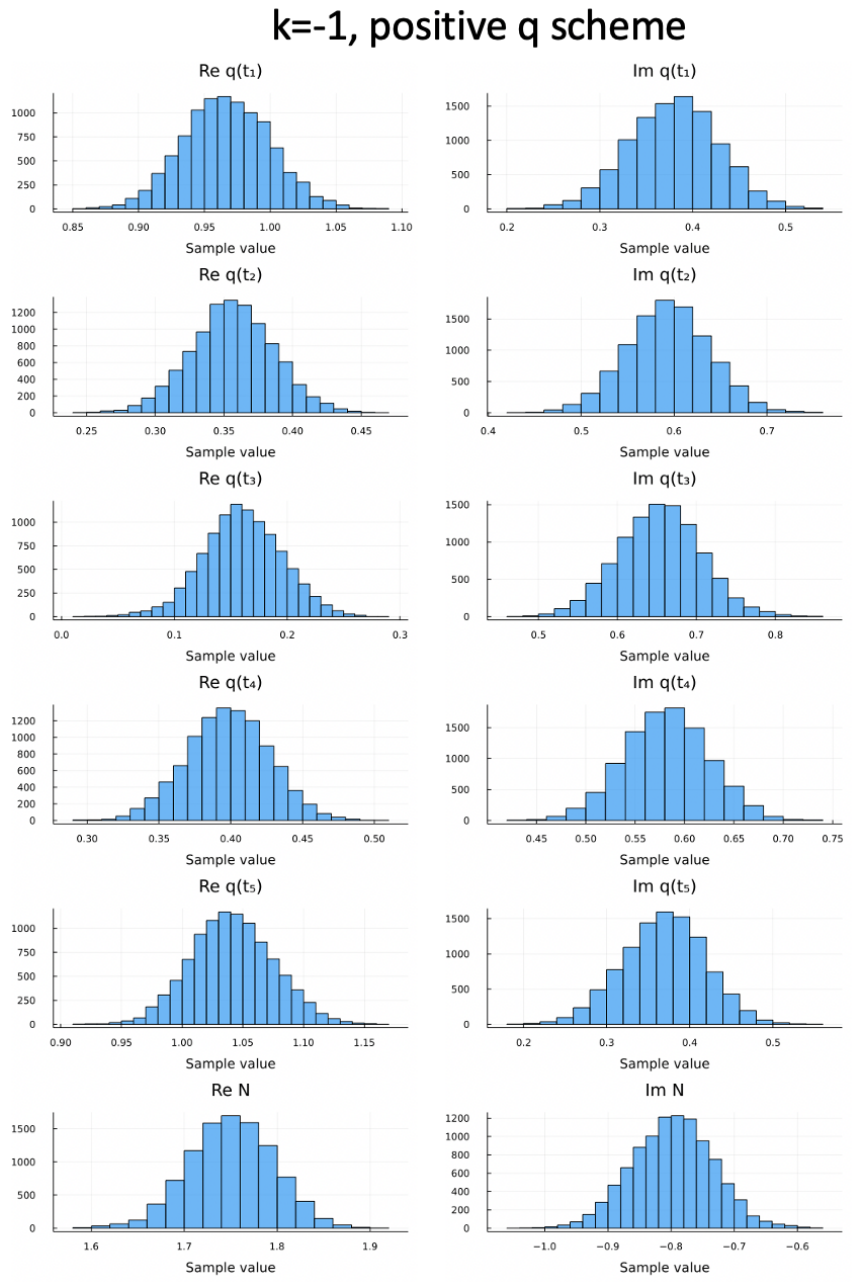

Figure 4.8: Histograms for the Monte Carlo sampling data for $k = -1$ in the positive q scheme. To reduce complexity the length of the samples is reduced from 1000 million to 1 million by sequentially picking the first element from every 1000 samples.

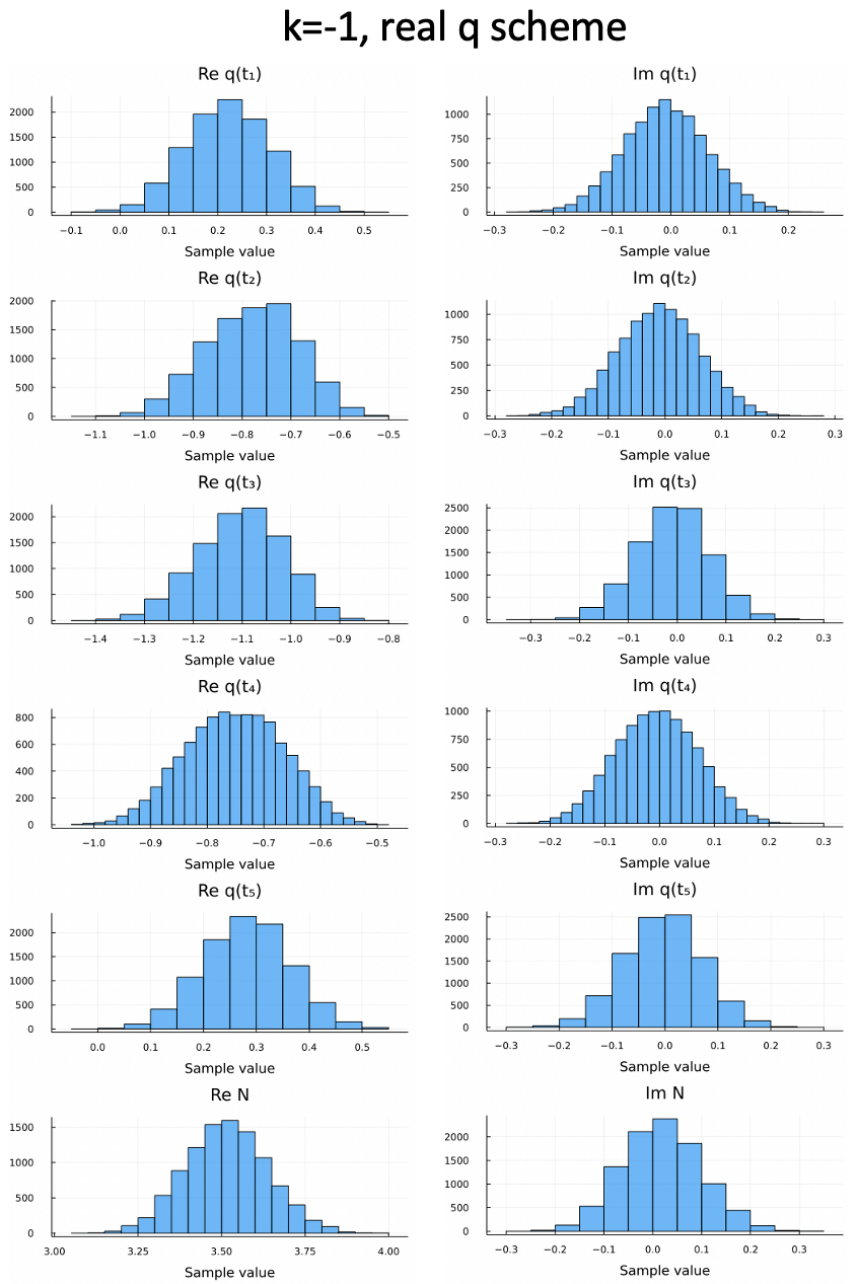

Figure 4.9: Histograms for the Monte Carlo sampling data for $k = -1$ in the real q scheme. To reduce complexity the length of the samples is reduced from 1000 million to 1 million by sequentially picking the first element from every 1000 samples.

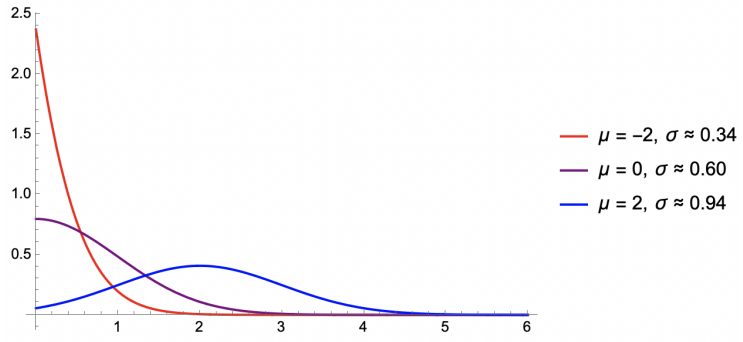

Figure 4.10: Sections of Gaussian distributions centered around μ and normalized over $\mathbb{R}^+$. The standard deviation σ is smaller for smaller μ.

the $k = 0$ case the saddle point value q does get very close to zero, while in the $k = 1$ case it does not, larger fluctuations in $\Delta\chi$ is expected.

4.6.5 Breakdown of saddle point approximation

In previous works of quantum cosmology, it is common to apply saddle point approximation to the path integrals. Here in the $k = -1$ example the saddle point failed at capturing the quantum expectation values of the truly Lorentzian path integral. This shows that the technique of saddle point semiclassical approximation does not enjoy universal validity for Lorentzian quantum cosmology, and must be used with caution.[3]

4.7 Singularity avoidance

4.7.1 Is singularity avoidance trivial?

Even when the expectation values are close (as in the $k = 0$ case), the positive and real q schemes still differ on the critical issue of singularity avoidance. There are many non-trivial ideas of singularities avoidance in quantum gravity. For example, through discreteness, nonlocality, higher-order terms in the action, final boundary condition choices etc. However, there is a trivial alternative. A gravitational path integral may simply not include singular spacetimes in its sum [8].

For the minisuperspace model studied here, this indeed follows from including only Lorentzian configurations in the path integral. At the $q = 0$ singularity, the metric (4.3) is of signature $(\infty, 0, 0, 0)$. This is not of the Lorentzian signature, so it is automatically avoided in the truly Lorentzian path integral. In this sense, singularity avoidance is trivially achieved.

[3] When naive saddle point approximation on unrestricted domain does break down, it may be interesting to develop new saddle point approximation method on bounded domains along the line of, e.g., [147].

4.7.2 Tunnelling and no-boundary proposals

Interestingly, insisting on a strictly Lorentzian path integral for all time including $t = 0$ invalidates from the outset Lorentzian variants of the tunnelling/no-boundary proposals that set $q(0) = 0$. Therefore one must choose from: (1) allowing non-Lorentzian configurations in the path integral; or (2) rejecting boundary conditions that set q to zero.

Choice (1) calls for some additional specifications. Suppose quantum cosmology is governed by some fundamental theory of quantum gravity. Then how exactly are non-Lorentzian configurations included in the path integral for this fundamental theory? Is non-Lorentzianess only allowed at certain places of quantum spacetime but not others? If so, at exactly which kind of places, and why not at other places? One possibility is to consider non-Lorentzian pieces at the boundary of superspace, and allow this kind of non-Lorentzianness in the path integral [148]. However, this proposal needed to divide the boundary of superspace into regular and singular parts, and append additional rules associated with probability fluxes to these two parts. However, as far as we know the exact definition of the regular and singular parts of the boundary has never been written down in general [149], and this proposal still remains as an incomplete idea.

Choice (2) is dynamically less ambiguous since no additional rule is needed on how to include non-Lorentzian configurations. However, it leaves open the question of boundary conditions which can only be determined by other means. One possibility is to impose an ordinary Lorentzian boundary condition concentrated around small positive values of q. Another possibility is to give up on boundary conditions at small sizes of the universe and investigate boundary conditions for bouncing cosmology [150].

4.8 Discussions

Quantum cosmology based on Lorentzian path integrals is a promising avenue. However, many previous studies integrate the squared scale factor over the whole real line. This step introduces non-Lorentzian configurations into the path integral. Instead, a truly Lorentzian path integral should only include positive squared scale factor.

Here we studied and compared minisuperspace path integrals with real and positive squared scale factors. By restricting to Lorentzian configurations, the truly Lorentzian case enables the study of causal horizons and their quantum fluctuations, and achieves singularity avoidance by excluding singular minisuperspace geometries as non-Lorentzian. In addition, we find that the expectation values can differ much between the two cases. This happens in particular when the saddle point configuration does not belong to the strictly Lorentzian integration contour and is not connected to the strictly Lorentzian integration contour by the holomorphic gradient flow.

These results challenge the universal validity of saddle point approximation widely used in quantum cosmology. In particular this affects topics such as Lorentzian variations of tunnelling/no-boundary proposals, and the quantum completeness of inflation [123]. In these cases, the saddle point gets close to or below zero, so that it does not belong to the strictly Lorentzian integration contour. Instead of using saddle point approximation, a safer option is to compute the path integral directly. This can be done, for example, using the generalized thimble method adopted here.

We finish by a discussion on some topics to be understood better.

4.8.1 Negative q

Although the metric (4.3) has the $(-,+,+,+)$ signature only when $q > 0$, it has the $(+,-,-,-)$ signature when $q < 0$. One may wonder whether this rescues the real q scheme for a Lorentzian path integral, since the $(+,-,-,-)$ signature might also be considered Lorentzian.

However, an attempt at rescue face some outstanding issues, because the real q scheme path integral includes configurations where $q < 0$ at certain times and $q > 0$ at other times.

First, such a configuration involves signature change. It does not qualify as a Lorentzian spacetime in the usual sense such that the spacetime stays within the $(-,+,+,+)$ or the $(+,-,-,-)$ signature.

Second, in connecting the $q < 0$ and $q > 0$ parts of the configuration q has to cross 0. Here the metric has signature $(\infty, 0, 0, 0)$. This is not Lorentzian.

Third, when q crosses 0, it is not *a priori* clear what the causal structure is for that spacetime. Some additional rules are required to tell how causal paths travel across the singularity at $q = 0$. Without such a rule, the causal relation between two events from the $q < 0$ and $q > 0$ parts of spacetime remains unclear.

4.8.2 Inhomogeneity, anisotropy, and matter coupling

The present study is restricted to minisuperspace models. For further research it is certainly interesting to accommodate inhomogeneity and/or anisotropy in the truly Lorentzian setting. For example, the Bianchi types I and III, and Kantowski-Sachs models studied in [107], and the biaxial Bianchi IX model studied in [119, 120] may be simple enough as starting points to incorporate anisotropy. In a general non-perturbative setting, simplicial manifold models provide a systematic way to incorporate inhomogeneity and anisotropy in quantum cosmology [72, 104, 56, 63, 64, 65, 66, 67, 68, 69, 70]. Traditionally, simplicial quantum gravity is studied with respect to an Euclidean contour or an *ad hoc* complex contour, but there has been growing attention towards the Lorentzian case [23, 24, 5, 2, 3, 6, 8, 27, 93, 151]. In particular, the generalized thimble method employed here and in [2] may be applicable in studies of inhomogeneity and anisotropy.

Certainly one should also consider matter coupling in further works. In addition to coupling to scalar fields and investigate the inflation scenario, we also find alternative scenarios without inflation worth investigating [58, 152, 153, 154].

4.8.3 Lightcone topics

Another topic about simplicial models of direct relevance is irregular lightcone structures. In simplicial models, there is the question whether the path integral should include simplicial geometries with interior points attached to more or fewer than two lightcones. In [93, 27] this question is studied based on a comparison with the continuum minisuperspace model in the real q scheme. It is worth revisiting this topic given that the positive q scheme may yield a different result.

In Section 4.6, we noted that the $k = 0$ case exhibit larger causal horizon fluctuations than the $k = 1$ and $k = -1$ cases. For the $k = -1$ case, that the fluctuations are smaller is related to the

breakdown of saddle point approximation based on Einstein's equations. Whether this and other effects of lightcone fluctuations lead to any observable signatures is worth investigating further.

4.8.4 Singularity

Much of quantum cosmology is driven by the hope to understand singularities. In the recent wave of interest for Lorentzian quantum cosmology, the question has been raised whether singular geometries with $q = 0$ should be avoided in the path integral [126], as such geometries enter critical discussions about boundary conditions [118, 122, 125, 126, 135] and inflation [127]. In Section 4.7 we showed that in a strictly Lorentzian path integral, singularities are automatically excluded as non-Lorentzian. How this affects the above topics is an open question.

There are many attempts to find effective regular spacetimes that replace spacetimes with cosmological and black hole singularities. Some of these derive regular solutions from equations of motion of modified actions. The $k = -1$ example studied here shows that an effective singularity-free geometry that characterizes the quantum theory at leading order (e.g., gives the correct expectation values) need not obey the equation of motion from an action principle. It remains to be clarified how such alternative views on singularity avoidance stand to each other.

4.8.5 Analytic insights

In computing the oscillating complex path integrals, we applied the generalized thimble method [33] to overcome the numerical sign problem. This method would not have been available a decade ago. However, new methods for evaluating complex path integrals are being developed at a promising pace in the recent decade (see e.g., [31, 78, 79, 155] and references therein). We expect such technical tools to boost the study beyond semiclassical analysis for the Lorentzian path integrals.

That said, it is still beneficial to find analytic methods to complement the numerical methods. One idea is to identify the value of N so that the N-dependent saddle point $\bar{q}$ of (4.7) just falls within the Lorentzian domain. Flowing this pair of N-$\bar{q}$ values under the holomorphic gradient flow to the steepest descent contour may yield a close guess at the expectation values.

Chapter 5

Is singularity avoidance trivial?

Many non-trivial ideas have been proposed to avoid singularities in quantum gravity. In this short note I argue that singularity avoidance can be trivial in gravitational path integrals, because geodesically incomplete singular spacetimes are usually not included in the sum. For theories where this holds, there is no need to develop non-trivial ideas on singularity avoidance. Instead, efforts should better be directed to understand tunneling processes and complex-valued spacetimes.

5.1 Is singularity avoidance non-trivial?

Spacetime singularities mark the breakdown of General Relativity and form a major motivation for quantum gravity. Many ingenious ideas have been considered before to resolve singularities in quantum gravity. In some proposals it is important to make a judicious choice of variable, e.g., using loop variables as opposed to the metric variable [156, 157]. In some proposals it is important to choose the right action, e.g., add higher order terms to the Einstein-Hilbert action [158]. In some proposals it is important to pick the right boundary condition, e.g., impose a special final boundary condition at the singularities [159]. In some proposals it is important to understand singular solutions at a detailed level, e.g., quantize starting from BKL type solutions [160]. In some proposals it is important to note some nonlocal property of quantum gravity, e.g., in the collapsing shell models of [161, 162, 163]. These and other proposals leave the impression that singularity avoidance in quantum gravity is non-trivial and requires some ingenious input.

As a counterpoint this short note presents a trivial idea on singularity avoidance. I argue that generically spacetime configurations singular in the sense of geodesic incompleteness simply do not belong to gravitational path integrals. For theories where this holds, singularities are avoided trivially, and ingenious endeavours can be saved. Instead, efforts should be directed to understanding tunneling processes and complex solutions, as I shall explain below.

5.2 Singular configurations fall out

In a matter path integral on a classical spacetime, we sum only over matter configurations which assumes values everywhere in the region that the path integral refers to (Figure 5.1). It is natural to expect that the same holds for gravitational path integrals. Yet a geodesically incomplete

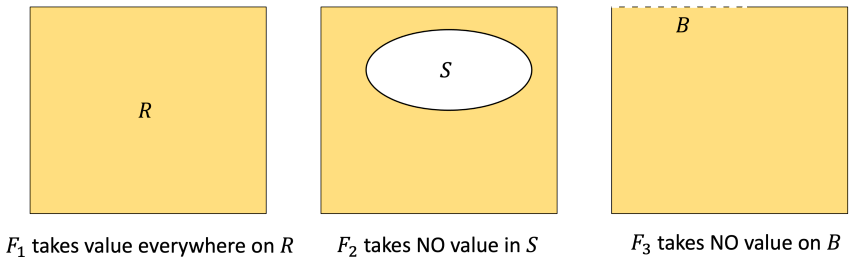

Figure 5.1: Matter configurations F_2 and F_3 do not assume values in the interior or on the boundary of the region R, so only F_1, which assumes value everywhere, belongs to the matter path integral on that region.

singular configuration does not meet this requirement. For such a configuration we could follow a causal curve for some finite affine parameter and find out that the gravitational variable stops possessing values. If gravitational path integrals are similar to the matter ones in this respect, singular configurations do not belong to the sum. Consequently the singular spacetimes relevant to the classical theory according to the singularity theorems [164] become irrelevant to the quantum theories defined by path integrals.

This plausibility argument may not hold for every theory of quantum gravity, but can be checked to hold in some explicitly defined gravitational path integrals. For example, causal dynamical triangulation [28], locally causal dynamical triangulation [165, 166] and Lorentzian simplicial quantum gravity [2] with a lightcone constraint [3] are defined by path integrals over piecewise flat spacetime configurations. In all these theories, any point in the interior of a configuration has two lightcones attached to it. A causal path reaching an interior point from one lightcone can always be extended away through the other lightcone. Therefore a causal path either extends indefinitely, or terminates when it reaches the boundary of the simplicial manifold if there is one.[1] To the extent that singular configurations are characterized by the inextendability of causal paths, they are simply not included in the path integrals.[2][3] Note that introducing matter to these theories does not change the conclusion, since the set of gravitational configurations of the path integral stays the same. In considering mixed boundary conditions, one needs to turn the path integral for amplitudes into "double path integrals" for probability weights [172, 173, 174]. Doubling up the path integral still does not change the conclusion about the absence of singular configurations in these theories, since the set of gravitational configurations in each individual copy of the path integral stays the same.

[1] In the latter case the gravitational variables (edge squared lengths) assume values on the boundary, so the boundary cannot be interpreted as a singularity.

[2] In variants of the above theories, a point of a spacetime configuration is allowed to have more or fewer than two lightcones [167, 168, 71]. If a point has just one lightcone, a causal path may not be extendable beyond it. This is usually interpreted as due to topology change rather than singularities to be avoided.

[3] In recent years, there has been an interest in incorporating topology changes in gravitational path integrals to match Bekenstein-Hawking black hole entropy (see [169] and references therein). In my view, one should stay cautious about these works, because they typically assume that all stationary points contribute to the path integral at the semiclassical level, and there is no fundamental justification to it. If one takes the view that the stationary points contribute because they lie on the original or deformed integration contour, then one should expect that only a subset of stationary points contribute [60]. Moreover, the central motivation for these works is the "central dogma" [170], which a non-dogmatic person may reasonably not subscribe to [171]. In fact, another way to read these works is that the potential falsity of the "central dogma" goes hand-in-hand with the potential falsity of uncontrolled topology changes.

Similar checks can be performed on other explicitly defined gravitational path integrals. While it is possible to encounter peculiar cases where singular configurations are included[4], the opposite is expected generically for a simple reason. We usually define

$$Z = \int DgA[g] \tag{5.1}$$

and the matter-gravity coupled integral $Z = \int DgD\phi A[g, \phi]$ by specifying how to sum over the values of a gravitational variable g on a lattice or some other structures. The case that g does not assume any value somewhere is excluded in this step.

5.3 Minisuperspace models

Besides fully general gravitational path integrals, minisuperspace path integrals offer a simplified setting for investigating singularity avoidance. In Suen and Young's proposal for quantum cosmology, the wave function for the universe is given by

$$\psi(h_{ij}, \phi) = \int_L DgD\phi \; e^{iS[g,\phi]}, \tag{5.2}$$

where "the set L includes all nonsingular Lorentzian four-geometries which induce [spatial geometry] h_{ij} on one of their boundaries, together with all [matter] fields ϕ regular on them" [114]. In de Sitter minisuperspace for pure gravity, it is found that the wave function tends to zero as the scale factor tends to zero [114].

Singularity avoidance of Lorentzian minisuperspace path integrals have also been investigated recently by the current author in [4]. The study emphasizes that the metric signature changes where the squared scale factor reaches zero for cosmological singularities, leading to the direct exclusion of spacetimes with such singularities in a strictly Lorentzian path integral. As a result, lightlike geodesics always extend from the past boundary to the future boundary without termination in all path integral configurations. As an application, for lightlike geodesics emanating from a fixed starting point on the past boundary, the quantum fluctuations of the exiting location on the future boundary is computed through the path integral. This provides a quantification of causal horizon fluctuations driven by quantum gravity.

In the above models, the strictly Lorentzian path integral sum exclude singular geometries by definition, and singularity avoidance is achieved trivially. The situation is more subtle in the path integrals for the no-boundary [44, 177] and tunnelling proposals [142] for quantum cosmology.

The no-boundary proposal is directly motivated by singularity avoidance. As presented by Hawking [178]: "Once one allows that singular histories could take part in the path integral, they could occur anywhere and predictability would disappear completely. [...] To implement the idea that the laws of physics hold everywhere, one should take the path integral only over nonsingular metrics." According to the no-boundary proposal, "the path integral for quantum gravity should be taken over all compact Euclidean metrics" [178] to achieve singularity avoidance: While in a Lorentzian metric, a singularity occurs when the spatial metric determinant h vanishes, this is not

[4]One example is Witten's original Chern-Simons formulation 2+1D quantum gravity, which includes non-invertible spacetimes where the vierbein variable is allowed to vanish [175]. Interestingly, in a more recent study Witten changed his mind to advance the view that such non-invertible spacetimes should be excluded [176].

necessarily the case in Euclidean geometry, where a vanishing 3-sphere can result from slicing a regular 4-geometry. However, the Euclidean gravitational path integral diverges due to conformal instability [20], and complex contours proposed to sidestep the issue not only lack justification from fundamental theories, but also face ambiguities in the choice of contour [109]. One may therefore reasonably refrain from regarding the no-boundary path integral as a realization of trivial singularity avoidance.

Vilenkin's tunnelling proposal posits that the universe spontaneously emerges through quantum tunnelling from "nothing" [43]. In minisuperspace path integral realizations of the idea, the initial scale factor is set to zero [142], resulting in a cosmological singularity. Sometimes this is characterized as a "nonsingular boundary condition", because in the terminology of [148], the boundary of superspace is divided into two parts: the "nonsingular boundary", encompassing singular 3-geometries resulting from slicing regular 4-geometries, and the "singular boundary", containing the remaining 3-geometries. Although a vanishing scale factor belongs to the nonsingular boundary of homogeneous and isotropic minisuperspace in this specific terminology, it realizes a cosmological singularity in the usual sense. Therefore the tunnelling path integral should count as a counterexample rather than an example for trivial singularity avoidance.

5.4 Classical approximations

If singular spacetimes do not even arise in a generic gravitational path integral, how come that they are essential in studies of black hole and cosmology in classical gravity?

In a classical theory, what we care about is a differential equation – the classical equation of motion. Singular spacetimes solve the differential equation for certain initial conditions, so are relevant. This is to be contrasted with quantum theory, where the path integral we care about is an integral. Instead of solving any differential equation, the relevant configurations are enumerated according to the values the gravitational variables take as explained above. This difference makes singular spacetimes irrelevant.

On the other hand, the path integral usually receives dominating contributions from its stationary points, which are solutions to the Euler-Lagrange equation. Will singular spacetimes not become relevant in the leading-order approximation in light of this, even for the quantum theory?

The answer is no. What singular spacetimes solve are initial value problems. Yet the stationary points need to solve boundary value problems [36]. Therefore singular spacetimes are still irrelevant for the leading approximation of the quantum theory.

To see this in more details, consider a path integral of the form

$$Z[h] = \int_h DgA[g], \tag{5.3}$$

where the sum obeys the condition h on the boundary B of the region. A special case of common interest is when $Z[h]$ gives the transition amplitude between h_1 and h_2, the two parts of h on the two components B_1 and B_2 of B. In general, the stationary points $\overline{g}$ for (5.3) need to solve the equation of motion for the boundary condition h, i.e., $\overline{g}|_B = h$. This condition cannot be met by singular spacetimes, because they cannot take the value h on its total boundary. For instance, suppose h decomposes into h_1 and h_2, where h_1 coincides with the initial boundary geometry for a portion of a Schwarzchild black hole. A singular spacetime may solve the equation of motion

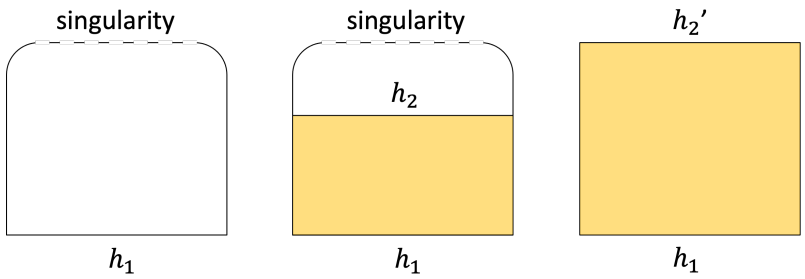

Figure 5.2: (a) A singular configuration with initial condition h_1 cannot match any final condition h_2. To also match some h_2, the configuration needs either to be truncated to remove the singularity, as in (b), or replaced by an entirely new configuration that shares h_1, as in (c).

under the initial condition h_1 alone, but it can never match h_2 on the other part of the boundary (Figure 5.2).

We may be interested in path integrals of a more general form

$$Z[\psi] = \int Dh \int_h DgA[g]\psi[h]. \tag{5.4}$$

A special case of interest is when h decomposes into h_1 and h_2 as above, and we consider

$$Z[\psi] = \int Dh_2\ \phi[h_2] = \int Dh_2 \int Dh_1 \int_{h_1 \cup h_2} DgA[g]\psi[h_1], \tag{5.5}$$

where the complex function $\psi[h_1 \cup h_2] = \psi[h_1]$ depends only on h_1. Here $\phi[h_2]$ is analogous to the wave function $\phi(x_2) = \int Dx_1 \int_{x(t_1)=x_1}^{x(t_2)=x_2} Dx\ A[x]\psi(x_1)$ of a non-relativistic particle at time t_2 given its wave function $\psi(x_1)$ at time t_1. Another special case of interest is when $\psi(h) = 1$ identically which yields Z as the full partition function. In general, the stationary points $\overline{g}$ for (5.4) are obtained by first solving $\delta_h(A[g]\psi[h]) = 0$ for the stationary points h_S on the boundary, and then solving $\delta_g(A[g]\psi[h]) = 0$ for $\overline{g}$ under each boundary condition h_S. In cases such as $\psi[h_1 \cup h_2] = \psi[h_1]$ when the boundary equation $\delta_h(A[g]\psi[h]) = 0$ is not very constraining, there could by many stationary points h_S. Yet $\overline{g}$ are obtained case by case, and in any case we solve $\delta_g(A[g]\psi[h]) = 0$ under a fixed boundary condition h_S. As explained in the last paragraph, singular spacetimes cannot be solutions.

Finally, coupling matter yields path integrals of the form

$$Z[\psi] = \int DhD\chi \int_{h,\chi} DgD\phi A[g,\phi]\psi[h,\chi], \tag{5.6}$$

where χ are the boundary matter configurations. For the same reasons as in the pure gravity case, geodesically incomplete singular configurations do not solve the Euler-Lagrange equation to form stationary points.

5.5 Tunneling and complex spacetimes

Singular solutions are essential in classical gravity for understanding black holes and cosmology. If singular spacetimes become irrelevant in the quantum theory and in its classical approximation, what replaces them for understanding black holes and cosmology?

A short answer is that they are replaced by tunneling processes whose stationary points are complex-valued spacetimes that satisfy the boundary conditions [60, 59, 179, 62]. These spacetimes solve the classical equation of motion without being singular, as the singularity theorems are evaded by allowing the spacetimes to be complex-valued.

The situation is quite similar to particle tunneling. In the classical theory energy conservation prevents the particle trajectory from extending beyond a potential barrier. However, in the quantum path integral the boundary conditions for detecting the particle beyond the potential barrier yield positive probabilities, accompanied by complex-valued solutions to the classical equation of motion [36, 37, 38]. The ban of energy conservation is evaded, because complex trajectories come with complex momentum so that the kinetic energy can be negative under the potential barrier.

While the basic scheme of this tunneling picture for singularity avoidance is clear, some open problems remain to be clarified. Firstly, how should we understand the complex spacetimes that arise in the classical approximations to singularity-resolving processes, when the original path integral includes only real-valued spacetimes in the sum? Mathematically, the complex stationary points can be understood through Picard-Lefschetz theory as belonging to the deformed integration contours in the complexified domain [60]. Yet physically, an intuitive understanding of the complex spacetimes seems to be missing. In particular, does a complex spacetime come with a causal structure? Are there lightcones on it? How do matter and information propagate in it? Secondly, the above qualitative picture holds for generic boundary conditions. Yet which specific boundary conditions should we use for the black holes and the cosmos of our world? Is this purely an empirical question, or are there principles that guide the choices? Thirdly, naive saddle point approximations valid for path integrals on unbounded domains may not be valid on bounded Lorentzian domains [4], and new techniques are needed for direct evaluation [4] or saddle points approximation [147] for the latter path integrals on bounded domains. In a theory where the trivial idea of singularity avoidance applies, these constitute some interesting questions to be investigated further.

5.6 Comments on some alternative views

I have presented a view that singularity avoidance is trivial in gravitational path integrals because singular spacetime configurations are generically absent in the sum. Given an explicitly defined gravitational path integral one could check if this view applies. When it does, the interesting tasks are to understand gravitational tunneling processes and their boundary conditions, as well as the physical meaning of the corresponding stationary points which are complex-valued.

From this perspective, current attempts at using *real-valued* regular spacetimes to give effective descriptions of singularity-resolving processes along the lines of [49] could be misguided. Should an individual spacetime be chosen to capture the essence of the quantum processes, a complex-valued spacetime is more appropriate than a real-valued one. The phenomenological consequences of complex-valued spacetimes for gravitational waves and black hole images are certainly also worth studying.

The trivial idea of singularity avoidance is also worth discussing in view of some contrary statements present in the literature. For example, in a relatively recent work [158] it is stated that in a gravitational path integral based on the metric variable g_{ab},

> all possible metric configurations (modulo diffeomorphisms) are being summed over. Thus, the singular spacetime metrics that constitute solutions of the field equations in GR are included in the path integral.

The problem with metric-variable gravitational path integral is that it is unclear how to define it in a non-perturbative and Lorentzian setting. In quantum field theories, the standard way to specify a path integral non-perturbatively is through lattices [180]. This route leads back to theories such as those discussed in Section 5.2. For the theories considered there singular configurations are seen to not belong to the path integral. Alternatively one might consider functional renormalization group for an in-principle non-perturbative specification of an Euclidean gravitational path integral. However it is unclear how to take this approach over to the Lorentzian setting [181] which is suitable for discussing spacetime singularities.[5] Without being referred to a Lorentzian path integral with an explicitly defined measure, it is difficult to be convinced of the claim that singular spacetime configurations are included in the path integral.

As another example from recent works, in [184] it is stated as an "essential condition" for a solution to the "information loss paradox" that for an Euclidean gravitational path integral (EPI):

> There exist at least two histories, say h_1 and h_2, that contribute to EPI, where h_1 is an information-losing history while h_2 is an information-preserving history.

Here "information-losing history" means "the semi-classical history of an evaporating black hole in which the unitary evolution would be lost when the black hole has completely evaporated" [184], and from Figure 1 of that paper one might infer that an "information-losing history" is geodesically incomplete. The main idea of [184] is to understand information propagation of quantum black holes as tunneling processes in gravitational path integrals. It seems the main points of [184] depends on the presence of an information-preserving history h_2 rather than an information-losing history h_1, and for reasons discussed in Section 5.5 I remain hopeful that the program of [184] to understand black hole information topics through gravitational tunneling will succeed.

The present trivial idea also differs from non-trivial ideas such as those listed in Section 5.1. What accounts for the differences? Firstly, the trivial idea is based on gravitational path integrals, whereas ideas such as [156, 157, 160, 161, 162, 163] are based on Wheeler-DeWitt or Schrödinger equation type models. Secondly, the trivial idea refers to explicitly defined path integrals, whereas ideas such as [158, 159] refer only to formal expressions of path integrals. Without an explicitly specified measure it is impossible to check if singular configurations belong to the path integral sum. In view of the present work, this could be an important missing step that changes the conclusion on how singularities are avoided in quantum gravity.

[5] This should not come as a surprise. There are cases where we know that the Euclidean and Lorentzian gravitational path integrals are inequivalent, and hence cannot be related by something like a simple Wick rotation. Historically, the *raison d'etre* for considering the Lorentzian causal dynamical triangulation in the dynamical triangulation program is to cure the issue of degenerate geometries of the Euclidean theory – and indeed the Lorentzian theory turned out to be different [102]. In simplicial quantum gravity, the same conclusion is reach. See the discussion around equation (3.17) of [23]. References [182, 116, 183] contain additional reasons in broader contexts to not equate Euclidean and Lorentzian gravitational path integrals.

Part II

Experience

Chapter 6

Experience in quantum physics

A theory of everything should not only tell us the laws for matter, gravity, and possibly boundary condition for the universe. In addition, it should specify the relation between theory and experience. Here I argue for a minimal prescription in extracting empirical predictions from path integrals by showing that alternative prescriptions are unjustifiable. In this minimal prescription, the relative probability for one experience is obtained by summing over all configurations compatible with that experience, without any further restriction associated with other experiences of the same or other experiential beings. An application to Wigner's friend settings shows that quantum theory admits objective predictions for subjective experiences. Still, quantum theory differs from classical theory in offering individualized as opposed to collective accounts of experiences. This consideration of experience in fundamental theories issues several challenges to popular quantum interpretations, and points to the outstanding need for a theory of experience in understanding physical theories of everything.

6.1 Introduction

A physical theory of everything is supposed to tell us [185]:

1. The dynamical laws for matter and gravity.
2. The laws for the boundary condition of the universe, if there are such laws.
3. The relation between the theory and experience.

Task 1 is the focus of particle physics and quantum gravity. Task 2 is a main topic of quantum cosmology. Task 3 is, well, a big embarrassment of quantum physics.

Task 1 is fulfilled most straightforwardly by starting with our best theory for matter, the Standard Model, and extending its path integral to sum also over gravity (Figure 6.1). As a result, we get partition functions of the form

$$Z = \int Dq \, e^{\frac{i}{\hbar}S[q]}, \tag{6.1}$$

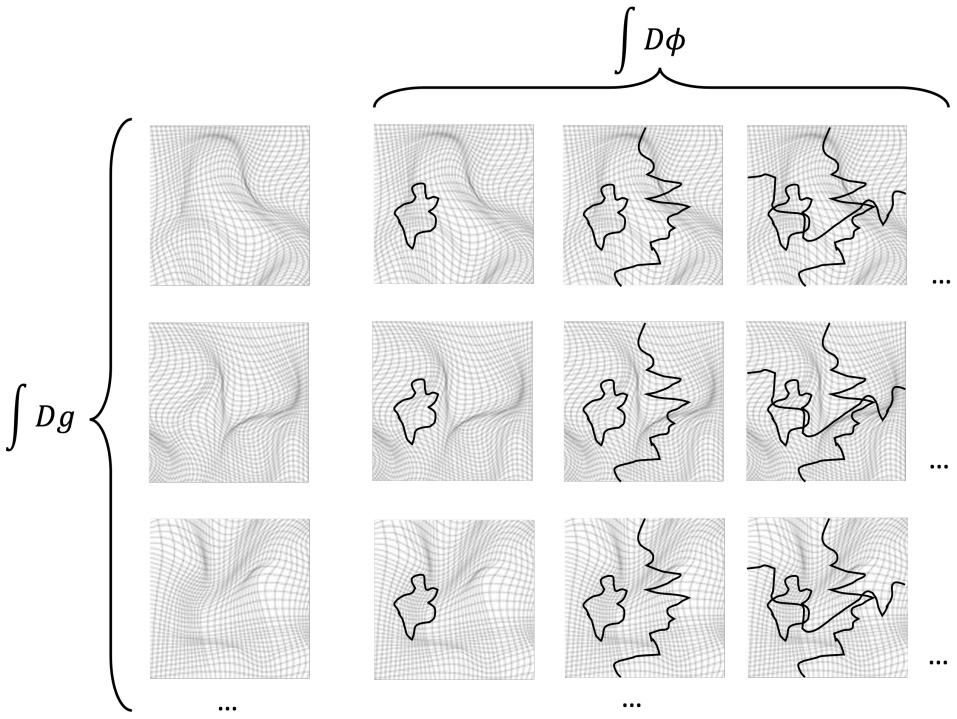

Figure 6.1: The path integral sum over everything $\int Dq = \int Dg \int D\phi$ contains gravity part $\int Dg$ and a matter part $\int D\phi$ (for simplicity, illustrated in the figure by particle configurations, instead of more realistic field or particle-string [1] configurations for the Standard Model).

where q contains both matter and gravity variables and $S[q]$ is the action. The matter part may be extended, e.g. to incorporate dark matter. The gravity part has more than one possible realization. See [186] for a survey of some popular approaches.

Task 2 invokes double path integrals of the form (see equation (4.8) of [187])

$$D[\rho] = \int Dq' \int Dq \; e^{\frac{i}{\hbar}(S[q]-S[q'])}\rho(q_b, q'_b), \tag{6.2}$$

where the path integral is doubled to account for the possible mixedness of the boundary condition ρ, which as an analog of the density operator takes as inputs the boundary configurations q_b and q'_b in the double copies. In case no law exists to fix the boundary condition, ρ is treated as an unknown parameter to be inferred empirically.

Task 3 is the focus of this work. This task is a big embarrassment, because we do not know exactly how the theory relates to experience quantum physics. From textbooks such as [13] we gather that in the path integral formalism, probabilities should be obtained through selecting a subset of configurations to integrate. However, there is no instruction on how to select in a theory of everything where all matter and gravity variables are subject to path integration. In particular, to predict Alice's experience at one moment, do we just select according to that experience alone?

Do we select on Alice's past and/or future experiences as well? Do we also select on Bob's and others' experiences, etc.?

These questions are important for cosmology, for instance in understanding how empirical probabilities are obtained in cosmology [188] and in resolving the issue of Boltzmann brain [189]. The questions are important for foundations of quantum field theory, for instance in understanding signalling constraints of measurements in QFT [190, 191, 192]. The questions are important for foundations of quantum physics, for instance in understanding Wigner's friend settings [193].

In this paper, I argue for a "minimal prescription" for extracting empirical predictions in theories of everything. The prescription is "minimal", because in predicting the probability for an experience, the path integral selects only for that experience. Alternative prescriptions which select for other experiences are systematically discussed, and are shown to be unjustifiable.

To illustrate the minimal prescription, I apply it to Wigner's friend [193] type thought experiments and show how touching base with fundamental theories fixes a unique and general prescription that accounts for the experiences of Wigner, Friend, and any other experiential beings. I explain why no commitment to any quantum interpretation is needed to arrive at the prescription, and why the recent Wigner's friend no-go theorems [194, 195, 196, 197, 198, 199] cannot help in arriving at the prescription.

Quantum theory with the minimal prescription gives an objective account for experiences, since everyone applies the same formula to account for anyone's experiences. The account is individualized, in the sense that a different formula is needed for each different experience. This is in contrast to classical theory, where a collective account of multiple experiences with the same formula is possible. In the world picture that emerges from the quantum theory of everything, matter and gravity path integral configurations coexist in superposition, whereas experience induces selection. Quantum states have no fundamental status, since the theory contains only boundary condition and experience selections, but not states that evolve and get updated in time. To the extent that experience is part of "reality", both boundary condition and experience selection describe "reality".

Based on these observations, I discuss some shortcomings of pilot-wave theories, collapse models, as well as Everttian, decoherent histories, relational, QBism, and neo-Copenhagen interpretations. Judged in the context of theory of everything, these approaches are in danger of being wrong, redundant, vague, and/or even superfluous. It remains to be seen if they offer any help in completing the remaining parts of three tasks for a theory of everything mentioned at the beginning.

As a note of scope, the discussion of theory of everything in this work is based in the path integral formalism.[1] Readers interested in canonical/algebraic formulations may find Don Page's "Sensible Quantum Mechanics" [200, 201, 202] relevant in accounting for experiences. Page's formalism is closely related to this paper in attributing quantum probabilities to individual experiences. However, it gives an essentially different prescription for making predictions in cosmology, where ambiguities associated with Boltzmann brain constitute a genuine problem [203, 204]. In contrast, the treatment presented here holds Boltzmann brain problem as a fake problem, as I will elaborate on elsewhere.

An outline of the paper can be found in the table of contents at the beginning.

[1] In my view, the most promising candidate theories of quantum gravity, e.g., Lorentzian simplicial quantum gravity [2], locally causal dynamical triangulation [165] etc., come from the path integral formalism.

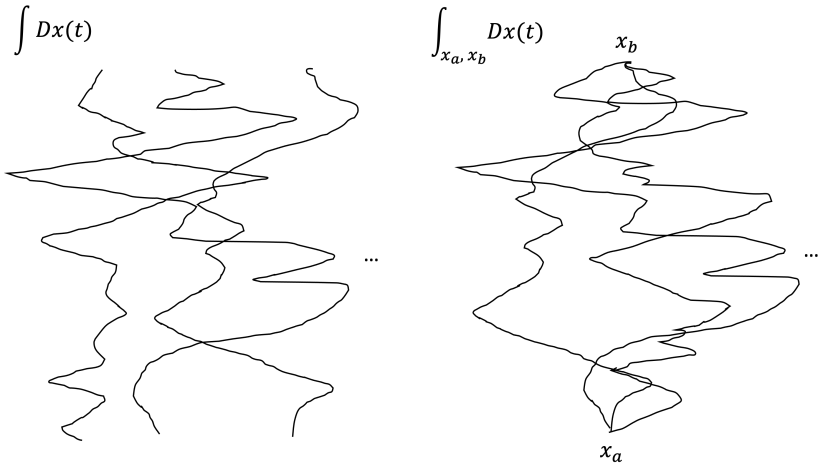

Figure 6.2: From the set of all particle paths (left), select those paths compatible with observational locations (right) to derive observational probabilities.

6.2 Experience in quantum physics

6.2.1 From particle to everything

How should we extract empirical predictions from a path integral? In chapters 1 and 2 of [13], Feynman and Hibbs offer a textbook treatment for a non-relativistic particle. To compute the probability for emitting and detecting a particle at a certain locations, one sums the amplitude over all paths originating and ending at these locations, and square it to obtain the probability (Figure 6.2). In other words, we select from all particle paths those compatible with the observational locations to obtain the probability for the observation.

In accounting for experiences in a theory of everything of the form (6.2), it is natural to also select the physical configurations compatible with the experience under consideration to derive its probability. However, as soon as we attempt to formalize this idea, questions arise (Figure 6.3). There are many experiential beings in the world, each going through many experiences. When extracting the probability for one experience of one being, should we select according to the experiences of all experiential beings, or just that one being? For that particular being, should we select according to multiple experiences, or just that one experience under consideration?

The answer is hard to find in previous works, even those that pay special attention to foundational questions of path integrals. The decoherent histories understanding of matter-gravity path integrals [187] suffers ambiguities in history selection, which makes it unclear how to extract empirical predictions [205]. The general boundary formalism [206] does not clearly fix a prescription on how many boundary conditions to impose, even when supplied with a relational interpretation [207] (In a universe containing many experiences and experiential beings, should we apply boundary condition to one experience? Some experiences? All experiences?). No other work I know of addresses the questions either. To proceed, we must find our own way.

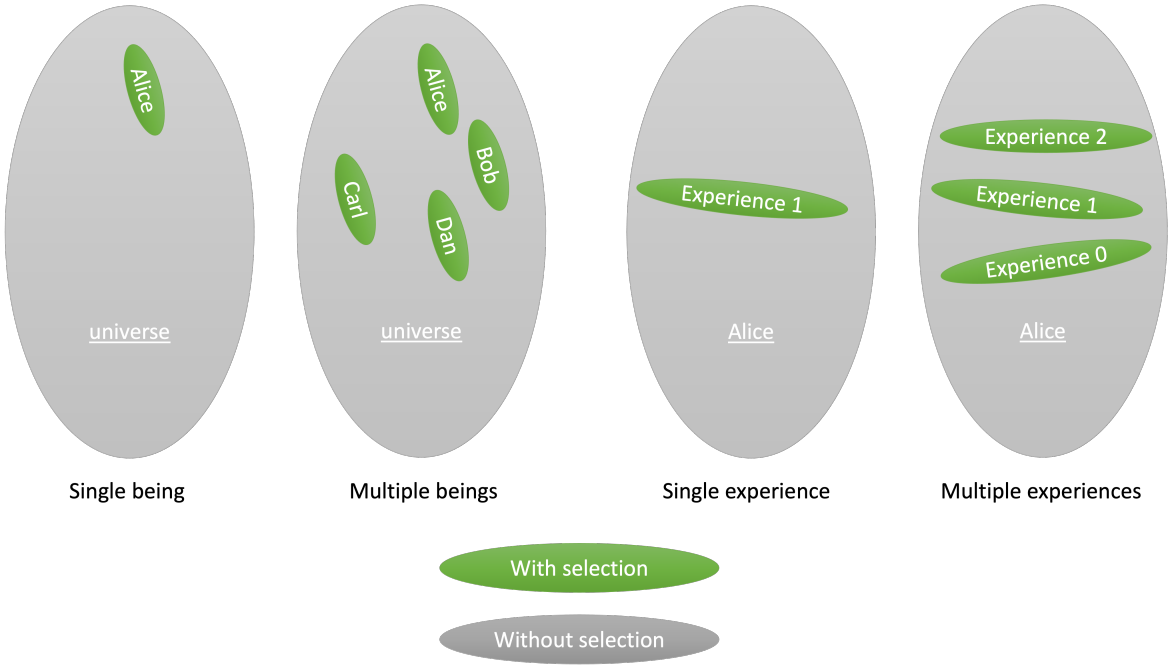

Figure 6.3: Different possibilities of selection for experiences.

6.2.2 General formula for experience

I argue below that to extract the relative probability for some experience, we should apply selection for just that one experience. Consider the empirical probabilistic predictions in the form of conditional probabilities

$$p(e_i|c). \tag{6.3}$$

Here c stands for a physical condition that enables a set of possible experiences $\{e_i\}_i$. For example, c could label the momentary (which may extend in time) physical configuration of a human being, incorporating all matters relevant for what he/she experiences next. In principle, this should fix the set $\{e_i\}_i$ of all possible next experiences.

In the **minimal prescription**, we select for a single experience for a single experiential being in the path integral. This corresponds to following the scenario of the first and third pictures of Figure 6.3. The probability for e_i to actualize under condition c is

$$p(e_i|c) = N \int_{e_i,c} Dq' \int_{e_i,c} Dq \; e^{\frac{i}{\hbar}(S[q]-S[q'])} \rho(q_b, q'_b). \tag{6.4}$$

The subscripts e_i, c indicate that in both branches of the double path integral, we select physical configurations compatible with the condition c and the experience e_i. The normalization constant N is fixed by requiring $\sum_i p(e_i|c) = 1$.

There are several open questions on how to implement the selection exactly. For instance, it is unclear what physical conditions enable experiences, what physical configurations correspond to experiences (this would constitute a definition of experience in terms of physics), what the list of all possible experiences exactly is under a given condition [208, 209]. It is unclear if human

experience supervenes on just the brain physical configurations, or if some part of the bodily configurations are also relevant [210]. It is unclear how much the physics of experience varies from species to species [211]. It is unclear at the mathematical level if the selections should always be sharp (a configuration is either included or excluded in the path integral), or can be unsharp (multiply the configurations by a weighting function taking non-binary values). ...

These are important scientific questions needing multidisciplinary inputs (see Section 6.6 for some further discussions). On the other hand, even without addressing these questions, just the form of the formula (6.4) without the details already has rich implications in physics. Specifically, (6.3) stands on three assumptions about experiences: (1) experiences are enumerable and distinguishable, i.e., it is meaningful to talk about e_i and e_j with $i \neq j$ as two different experiences; (2) experiences are enabled by some conditions, i.e., it is meaningful to talk about c as a condition that enables e_i; (3) experiences and their conditions have a physical basis, i.e., it is meaningful to relate $p(e_i|c)$ to a physical theory of everything. It is not impossible for one or more of these assumptions to fail, but they are quite reasonable to make to start with in understanding empirical predictions in physical theories of everything. Once these assumptions are made, we can accept that (6.3) can be related to physical theories of everything. Then Section 6.2.3 below offers an argument to accept (6.4), whose implications in physics will be explored in the rest of the work. Importantly, none of the steps require answers to the open questions about experience listed above.

6.2.3 Why select just one experience?

Before moving on to the implications, I should explain why we should select according to just one experience in the probability formula (6.4). To start with, note that:

Joint experience is not experience

Consider the combination of some experience of Alice with some experience of Bob, or with some other experience of Alice herself. The result is not an experience experienced by any being. Therefore to make probabilistic predictions for one experience, it is reasonable to consider selecting on that experience and that experience only.

In contrast, alternative selection scenarios can be shown to be quite unreasonable (Figure 6.4). If we were to select just for $n = 1$ experience, it is certainly reasonable to select according to one under consideration as in (6.4), instead of a different experience. If we were to select for $n > 1$ experiences, there is no n that is preferred for any reason. Therefore any particular choice of n lacks a good justification.

One might ask, why do we not select according to the number of experiences, or the number of experiential beings, that exist in the universe? The problem is that there is no fixed number of experiences or experiential beings in the path integral over everything (6.2). In any quantum region [206], the path integral sums over gravity and matter configurations with zero, one, two ... experiences, so it is meaningless to talk about a definite number of experiences or experiential beings that exist in the universe.[2]

[2]The boundary condition cannot help in selecting some particular n value or some set of n values of experiences that exist in the path integral configurations, because it only refers to the boundary part of the path integral configurations, but not the interior part where paths with different numbers of experiences are summed over.

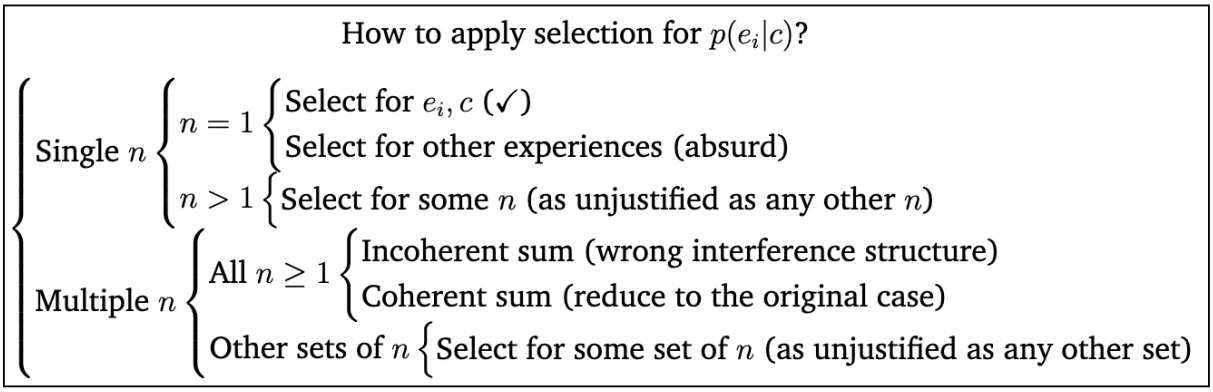

Figure 6.4: Structure of argument for picking the selection prescription for $p(e_i|c)$. Here n is the number of experiences to be selected for.

The above analysis indicates that $n = 1$ is the only reasonable option, if a single number of experiences is to be selected. An alternative is to select for a range of n values. In this case, the first option is to sum over all $n \geq 1$. Here $n = 0$ gets excluded since we want to make sure that the experience under consideration is selected. The sum could be performed either coherently or incoherently. A coherent sum over amplitudes for all numbers n, keeping the selection for the experience under consideration fixed in all the terms of the sum, actually gives back (6.4). On the other hand, an incoherent sum over probabilities for all numbers n is not a reasonable choice. This is because all quantum regions contain experiential configurations, so the result will exhibits a drastically different interference structure whose predictions will differ from empirical evidence. The second option is to sum over a different set of multiple n values. The problem, like in the case of selecting for a single $n > 1$, is that within quantum theory there is no particular set of n that is preferred for any reason. Therefore any choice lacks a good justification.

If one is willing to go beyond quantum theory, there could be mechanisms that fix one or a set of n values. For instance, the modified dynamics of consciousness collapse models [212, 213, 214, 215, 216] may fix some n value(s) through its stochastic reduction. In this paper, I focus on accounting for experience within quantum theory, so will not discuss such possibilities further (however, see Section 6.5 for some critical remarks on collapse models).

6.2.4 A toy model

Formula (6.4) is rather general and abstract. It helps to illustrate it in a concrete setting with three simplifying assumptions. (1) Consider a simplified setting where spacetime is reduced to classical. For instance, consider a saddle approximation to the gravitational path integral, where one classical solution to the gravitational boundary value problem dominates the path integral. (2) Moreover, assume that the matter degrees of freedom obey a unitary time evolution for some time foliation of the classical spacetime. (3) Assume further that the spacetime is sufficiently non-degenerate, so that we can locate experiences sequentially in time by selecting on the gravitational configurations.

Now consider a set of experiences $\{e_i\}_i$ possible under the condition c. Assume the latter to be implemented by the projector P at time t_1', and the former to be implemented by the projectors

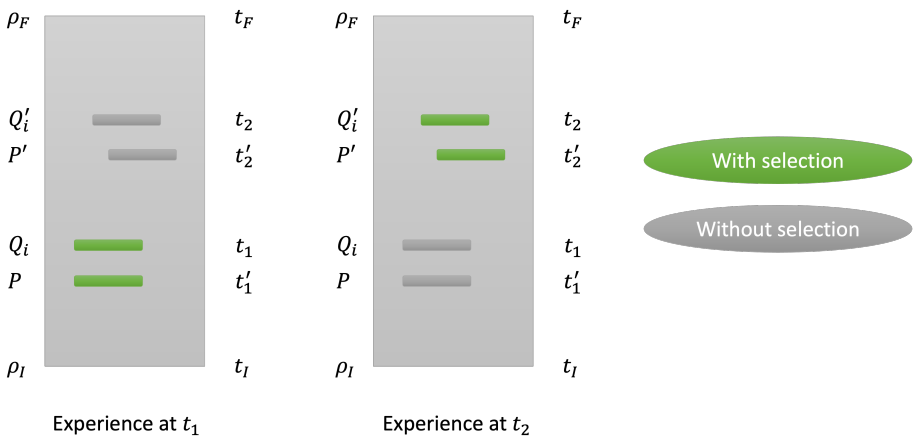

Figure 6.5: Different projections are activated for different experiences.

Q_i at time $t_1 > t_1'$ (Figure 6.5). Then

$$p(e_i|c) = N \operatorname{Tr}\Big[\rho_F U(t_F,t_1) Q_i U(t_1,t_1') P U(t_1',t_I) \rho_I U^\dagger(t_1',t_I) P U^\dagger(t_1,t_1') Q_i U^\dagger(t_F,t_1)\Big], \tag{6.5}$$

where N is fixed by setting $\sum_i p(e_i|c) = 1$, and as a further simplifying assumption, the boundary condition factorizes into ρ_I and ρ_F on the initial and final boundary of the universe. The unitaries $U(t_b,t_a)$ between times t_a and t_b is to be derived from the path integral propagators. In the Heisenberg picture where the operators

$$O(t) = U^\dagger(t,t_0) O U(t,t_0) \tag{6.6}$$

are time-dependent (t_0 is some reference time), the formula becomes

$$p(e_i|c) = N \operatorname{Tr}\big[\rho_F(t_F) Q_i(t_1) P(t_1') \rho_I(t_I) P(t_1') Q_i(t_1)\big], \tag{6.7}$$

where the initial and final boundary conditions lie at times t_I and t_F.

Suppose we want to model the experience of the same being with memory, at a later time $t_2 > t_1$ (Figure 6.5). Then we only impose selection for that experience to obtain

$$p(e_i'|c') = N' \operatorname{Tr}\big[\rho_F(t_F) Q_i'(t_2) P'(t_2') \rho_I(t_I) P'(t_2') Q_i'(t_2)\big], \tag{6.8}$$

where c' is a different condition with projector P' at $t_2' \leq t_2$, and $\{e_i'\}_i$ is a different set of experiences with projectors Q' at t_2.

The memory is encoded in P'. For instance, when the initial boundary condition is pure, i.e., when $\rho_I(t_I) = |\psi\rangle\langle\psi|$, the memory of experience e_j at t_1 and condition c at t_1' can be incorporated by setting $P'(t_2') = |\psi_j\rangle\langle\psi_j|$ in the Heisenberg picture. Here

$$|\psi_j\rangle = P''(t_2') Q_j(t_1) P(t_1') |\psi\rangle, \tag{6.9}$$

where $P''(t_2')$ implements further conditioning not already implemented by $Q_j(t_1)P(t_1')$. Plugging it in (6.8) yields

$$p(e_i'|c') \propto \operatorname{Tr}\big[\rho_F(t_F) Q_i'(t_2) P''(t_2') Q_j(t_1) P(t_1') \rho_I(t_I) P(t_1') Q_j(t_1) P''(t_2') Q_i'(t_2)\big]. \tag{6.10}$$

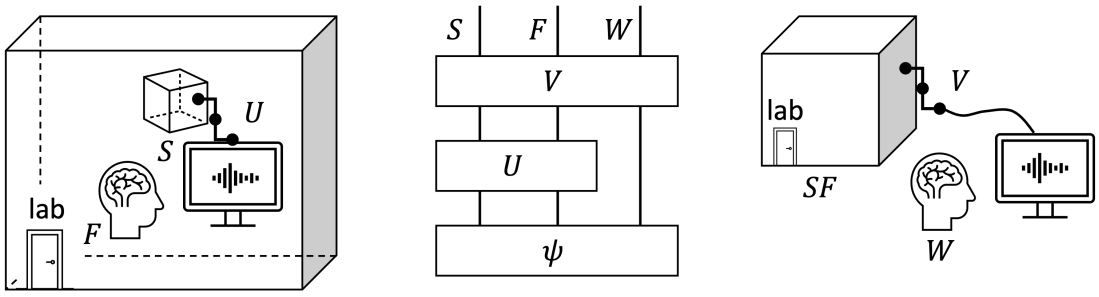

Figure 6.6: Wigner's friend setting. Left: Inside a lab, Friend (F) interacts with a system (S) through unitary evolution U. Right: Afterwards, Wigner (W) outside the lab interacts with system-Friend through unitary evolution V. Middle: circuit diagram for the interactions.

Even when selections are made only at t_2' and t_2, the influence of previous events are reflected in the current condition c' because of the "memories" encoded in $P'(t_2')$.

This toy model only gives a rough account for the experiences of experiential beings with memory. In a more realistic account, the condition encoding memory, P', should refer to the particular local bodily physical configurations of the experiential being, so will take a different from than that given by (6.9). This unrealistic aspect should be kept in mind when using the toy model.

6.3 Wigner's friend

Wigner's friend thought experiment [193] (Figure 6.6) poses the question whether in describing the experience of one being, other beings could be in superposition of physical configurations corresponding to different experiences. This setting provides an ideal ground to illustrate the prescription of the previous section, which answers the question with a resounding yes.

6.3.1 Setting

Consider the Wigner's friend setting of Figure 6.6 with the global unitary evolution

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|0\rangle_S + |1\rangle_S)\,|f\rangle_F\,|w\rangle_W \tag{6.11}$$

$$\xrightarrow{U} |\psi(1)\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{SF} + |11\rangle_{SF})\,|w\rangle_W \tag{6.12}$$

$$\xrightarrow{V} |\psi(2)\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{SF} + |11\rangle_{SF})\,|0\rangle_W\,. \tag{6.13}$$

At time $t = 0$, the system, describing for instance the spin of a particle, is initialized in the state $\frac{1}{\sqrt{2}}(|0\rangle_S + |1\rangle_S)$. Friend's and Wigner's physical configurations are initialized in some states $|f\rangle_F$ and $|w\rangle_W$.

At time $t = 1$, the global state evolves to $|\psi(1)\rangle$ under U, which couples Friend to the system. This constitutes the premeasurement unitary interaction for Friend to measure the system with the projections $\{|0\rangle\langle 0|\,, |1\rangle\langle 1|\}$.

At time $t = 2$, the global state evolves to $|\psi(2)\rangle$ under V, which couples Wigner to system-Friend. This constitutes the premeasurement unitary interaction for Wigner to measure system-Friend with the projections $\{|\Phi\rangle\langle\Phi|, \mathbb{1} - |\Phi\rangle\langle\Phi|\}$, where

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{SF} + |11\rangle_{SF}). \tag{6.14}$$

Without any experiential selections, the unitary evolution does not yield probabilities for any experience. The question is to decide on the projections to imply in accounting for the experiences of Wigner and Friend.

6.3.2 Wigner's experiences

Suppose in a simplified setting all assumptions leading to the toy model of Section 6.2.4 hold, so we can employ the model there. Let us suppose further that Wigner's experiences are characterized by projectors Q_i, and the condition only retains the most recent memory by applying the projection at the closest previous time (this amounts to implementing the condition by P' of (6.9) and setting the Heisenberg picture projectors P'' and P to identity in (6.9)). Suppose further that

$$\rho_F = I, \tag{6.15}$$

i.e., the final boundary condition is the identity in (6.5).

Then (6.7) yields, in the Schrödinger picture,

$$p(j|i) = N \operatorname{Tr}\left[Q_j U Q_i |\psi(0)\rangle\langle\psi(0)| Q_i U^\dagger Q_j\right], \quad t = 1, \tag{6.16}$$

$$p(j|i) = N \operatorname{Tr}\left[Q_j V Q_i |\psi(1)\rangle\langle\psi(1)| Q_i V^\dagger Q_j\right], \quad t = 2. \tag{6.17}$$

for experience j conditioned on memory for the previous experience i. For example, when there are only two possible experiences with projectors

$$\{Q_0 = \mathbb{1}_{SF} \otimes |0\rangle\langle 0|_W, Q_1 = \mathbb{1}_{SF} \otimes |1\rangle\langle 1|_W\}, \tag{6.18}$$

plugging in (6.16) yields that at time $t = 1$, $p(j|i) = \delta_{i,j}$, where the normalization factor N is fixed to be $|\langle i|w\rangle|^2 = 1$. This indicates that at $t = 1$ the experience must stay the same as previously.

At $t = 2$, the explicit form of V is needed to compute the result for (6.17). As an example, let W's state space be spanned by $\{|w\rangle, |w_\perp\rangle\}$ with

$$|i\rangle = \alpha_i |w\rangle + \beta_i |w_\perp\rangle, \quad k = 0, 1, \tag{6.19}$$

and F's state space be spanned by $\{|\Phi\rangle, |\Phi_1\rangle, |\Phi_2\rangle, |\Phi_3\rangle\}$ with $|\Phi\rangle$ given in (6.14). Suppose

$$\begin{aligned} V : &|\Phi w\rangle \mapsto |\Phi 0\rangle, |\Phi_k w\rangle \mapsto |\Phi_k 1\rangle, \\ &|\Phi w_\perp\rangle \mapsto |\Phi 1\rangle, |\Phi_k w_\perp\rangle \mapsto |\Phi_k 0\rangle, \quad k = 1, 2, 3. \end{aligned} \tag{6.20}$$

Then indeed V is compatible with (6.13). Plugging in (6.17) yields

$$p(j|i) = N \| \langle j| \, V (|w\rangle\langle w| + |w_\perp\rangle\langle w_\perp|) \, |\Phi i\rangle \, \langle i|w\rangle \|^2 \tag{6.21}$$

$$= N |\delta_{j,0} \alpha_i \alpha_i^* + \delta_{j,1} \beta_i \alpha_i^*|^2 \tag{6.22}$$

$$= |\delta_{j,0} \alpha_i + \delta_{j,1} \beta_i|^2, \tag{6.23}$$

where the normalization factor N has been fixed to be $|\alpha_i|^2$. This indicates that the probabilities for experience j at $t = 2$ are $|\alpha_i|^2$ for $j = 0$ and $|\beta_i|^2$ for $j = 1$.

6.3.3 Friend's experiences

The formulas for Friend's experiences are entirely analogous. Suppose Friend's experiences are characterized by projectors R_i (e.g., $\{R_0 = \mathbb{1}_{SW} \otimes |0\rangle\langle 0|_F, R_1 = \mathbb{1}_{SW} \otimes |1\rangle\langle 1|_F\}$). Then the probabilities take the same form as Wigner's (6.16) and (6.17):

$$p(j|i) = N \, \mathrm{Tr}\Big[R_j U R_i \, |\psi(0)\rangle\langle\psi(0)| \, R_i U^\dagger R_j\Big], \quad t = 1, \tag{6.24}$$

$$p(j|i) = N \, \mathrm{Tr}\Big[R_j V R_i \, |\psi(1)\rangle\langle\psi(1)| \, R_i V^\dagger R_j\Big], \quad t = 2. \tag{6.25}$$

Once explicit forms of U and V like (6.20) above are given, the probabilities can be computed explicitly.

More generally, there can be alternative settings with possibly more experiential beings interacting in arbitrarily complicated ways. In all these cases, the probability formulas would take the same form as (6.16), (6.17), (6.24) and (6.25). The prescription of Section 6.2 applies in generality.

6.3.4 Contrasting with previous treatments

The present treatment of the Wigner's friend setting within quantum theory contains some differences from some previous treatments.

In [217], DeBrota, Fuchs, and Schack give an analysis of Wigner's friend settings from a QBism perspective. On page 1869 of [217], one finds the statement: "the friend [...] amounts to assigning a quantum state to herself, which violates the QBist tenet that there must be a clear separation between agent and measured system". In contrast, the formulas in the present section do not forbid someone from assigning a state to herself. For example, when Friend makes a prediction for Wigner's experiences, like everyone else she should use the formulas of Section 6.3.2 which assign a state to herself. Conceptually, this highlights the **objective** nature of the quantum empirical predictions, which is easily forgotten in emphasizing the subjective nature of experiences [217].

In some treatments of Wigner's friend settings (e.g., [196, 218]), the humans of Wigner and Friend are replaceable by computers/automatic machines without making a difference. This is in contrast with the experiential treatment given here where experiential beings are special. In a world **without experiential beings**, there is no empirical prediction to make, and no selections to be applied to the path integral (and no projections to be applied to the state in the toy model). Therefore computers/machines without experience do not induce definite values for variables, and must be treated differently from humans with experience.

In [218], Baumann and Brukner imagines a "textbook treatment" of a Wigner's friend setting. In this treatment, even when predicting Wigner's observational outcomes, Friend applies projective state-updates for her own measurements. Baumann and Brukner point out that the "textbook treatment" differs from treatments by QBism, neo-Copenhagen, and relational interpretations within quantum theory, which leaves the impression that these **interpretations** make a genuine difference. In contrast, according to the analysis given here which touches base with **fundamental theories** with specified dynamical laws (encoded in the action for the path integral), no such interpretations are required to differ from the "textbook treatment". As argued in Section 6.2, formula (6.4) is the only reasonable prescription for extracting empirical predictions in the fundamental quantum theories considered here. This leads to the formulas of Section 6.3.2, which, in contrast to the "textbook treatment", apply projections only to Wigner when accounting for Wigner's experiences (by Friend, Wigner, or whoever else).

Finally, readers who know of the recent Wigner's friend no-go theorems [194, 195, 196, 197, 198, 199] may ask if these inform us anything on how to account for experiences within quantum theory. In particular, can the no-go theorems offer alternative arguments to rule out any prescription discussed in Section 6.2.3? The answer is no. The structure of argument in [194, 195, 197, 198, 199] is that quantum predictions violate certain Bell-type inequalities, so the assumptions leading to the inequalities cannot coexist if quantum theory is right. For this type of argument to take off, one must already know how to draw the correct quantum empirical predictions. However, knowing this would have fixed the prescription already, without considering any inequalities. Therefore the no-go theorems of [194, 195, 197, 198, 199] cannot help in picking out the correct prescription for quantum empirical predictions. The structure of argument in [196] is that certain assumptions put together yield a contradiction. If any prescription discussed in Section 6.2.3 obeys the assumptions, that would be ruled out by [196]. However, none of the assumptions applies, since the prescriptions only assign definite values to a physical variable when there is empirical selection (without selection, all values are summed over in the path integral), whereas the assumptions of [196] assume it is meaningful to talk about definite values of physical variable in themselves without regards to experience. Therefore the no-go theorem of [196] cannot help in picking out the correct prescription for quantum empirical predictions.[3]

This comparison with previous treatments reveals several conceptual questions that deserve to be clarified further. In which sense is the minimal prescription of Section 6.2 objective? In which is it not? What picture of Nature does quantum theory offer in the absence of experiential beings like us, as some "realists" tend to ask? In what ways do "fundamental" quantum theories differ from "non-fundamental" quantum theories? What is the role of quantum interpretations in accounting for experiences? I will address these questions in the rest of the paper.

6.4 Conceptual reflections

6.4.1 Objective vs. subjective

The empirical probabilities derived from (6.4) are **objective**, in the sense that everyone uses this same formula to account for anyone's experiences. For example, in the Wigner's friend setting,

[3] One may be tempted to ask, if the no-go theorems cannot rule out any prescription for quantum empirical prediction, what do they rule out at all? The one-paragraph discussion offered here does not address this question, but only explains why the no-go theorems cannot be used to rule out prescriptions discussed in Section 6.2.3. For in-depth critical discussions that do address the question, see, for instance, [219, 220] and references therein.

Wigner, Friend, and everyone else should use the formulas of Section 6.3.2 to account for Wigner's experiences.

Certainly the experiences themselves are subjective, in the sense that Wigner does not experience Friend's experience and *vice versa*. However, this does not change the objective nature of (6.4) in drawing empirical predictions.

An important consequence is that the quantum empirical probabilities admit an **objective propensity interpretation** (although this does not rule out other interpretations for the quantum empirical probabilities). Quantum interpretations such as QBism [221, 222, 223] hold that quantum probabilities are Bayesian probabilities for agents' beliefs. Here it is shown that even in view of extreme settings such as Wigner's friend, a Bayesian interpretation of quantum empirical probability is not a necessity.

6.4.2 Individualized vs. collective

There is a sense in which "objectivity" is weakened in (6.4) in comparison to classical theory. In classical theory, a **collective** account for multiple experiences is available. For instance, in a classical field theory, a field configuration throughout the spacetime is supposed to account for all experiences throughout this universe. In contrast, the quantum formula (6.4) offers only an **individualized** account, in the sense that a particular selection is used for each particular experience.

This individualized vs. collective dichotomy also distinguishes quantum theory based on (6.4) and collapse models. The distinction leads to falsifiable predictions. In the setting of Section 6.3.1, quantum theory based on (6.4) predicts that Wigner's experience is compatible with the measurement outcome of $|\Phi\rangle\langle\Phi|$ with (6.14), while collapse models rule out such coherent macroscopic superposition of Friend in accounting for Wigner's experience.

To people who object (6.4) on philosophical grounds for failing to give a collective account, or for being "soliptic", the response is that these objections are insignificant. A theory can be true or false (e.g., passing or failing the kind of empirical test discussed above) independent of whether or not it gives a collective account, and whether or not it is "soliptic".

6.4.3 Presence vs. absence

Some "realists" aspire to understand stable set of properties of matter in and of itself, without regard to human perceptions [224]. A path integral like (6.2) does suggest a very simple picture of what Nature is like, when no experiential beings are present. Without experiential beings, there is no experience selection. All path integral configurations for matter and gravity coexist in superposition.

The problem is that this picture, or any picture without experiential beings, is not empirically verifiable. It is not empirically verifiable because there is no experience without experiential beings. If someone gives a different picture in which the universe becomes a banana whenever experiential beings are absent, we cannot give empirical evidence to rule out the case. In Section 6.6 an alternative world picture is offered, which takes as a starting point the presence of experiential beings.

	Wrong	Redundant	Vague	Superfluous
Pilot-wave theories	✓			
Collapse models	✓			
Everettian	✓			✓
Decoherent histories		✓		✓
Relational		✓		✓
QBism			✓	✓
Neo-Copenhagen			✓	✓

Table 6.1: Dangers faced by interpretations, judged in the context of theory of everything

As an alternative, a "realist" may ascertain the presence of experiential beings, and aspire for a world picture that admits a collective account of experiences in the sense of Section 6.4.2. As explained there, quantum theory does not meet this hope. Whether this is good or bad should be decided by empirical tests, such as the kind mentioned in Section 6.4.2, which distinguish theories that come with collective accounts and theories that do not.

6.4.4 Boundary condition and experience selection vs. state

The account of Wigner's friend setting of Section 6.3 touches base with fundamental theories with specified dynamical laws encoded in the path integral action. In Section 6.2.4 an explicit explanation is given on how to start from formula (6.4) for the fundamental path integral to arrive at the formulas of Section 6.3 for the toy models. This explanation clarifies that quantum states evolving in time have no status in the fundamental path integral (6.4),[4] which only refers to the boundary condition of the universe and experience selections. States evolving in time arise only after imposing simplifying assumptions.

What is the nature of the quantum state? Does it describe our knowledge of reality, or reality itself? Much discussion in the literature treat these as essential questions for understanding quantum theory. Yet from the perspective of the fundamental path integral they are not essential questions, because the state is a dispensable concept.

What about the nature of the boundary condition and the experience selections? Do they describe our knowledge of reality, or reality itself? In (6.4), the boundary condition is kept fixed for all experiences, while the experience selection varies for different experiences. To the extent that experience is part of "reality", both are used in describing "reality". Neither describes our knowledge *per se*, although knowledge can be formed about the boundary condition and the experience selections.

6.5 On quantum interpretations

What does the analysis of experience in theories of everything inform us about quantum interpretations? In this section I explain why judged in this context of theory of everything, all the interpretations listed in Table 6.1 face major issues.

[4] See, e.g., [187, 225, 206] for related discussions.

6.5.1 Wrong?

To fulfill the first task for a theory of everything (Section 6.1), a putative theory should offer the correct laws for matter and gravity. The high danger in pilot-wave theories [226] and collapse models [227] is that they supply the wrong laws for matter. The high danger in Everettian interpretations is that they supply the wrong laws for gravity.

Pilot-wave theories and collapse models

For laws of matter, the best theory we have is the Standard Model, which passes stringent tests from particle physics experiments. The currently available pilot-wave theories and collapse models cannot correctly reproduce the Standard Model predictions [228, 229]. Future will tell if further developments can overcome this challenge.[5]

Everettian interpretations

The goal of Everett's original paper [230] was to present "a reformulation of quantum theory in a form believed suitable for application to general relativity". Back in 1957, far less is known about quantum gravity, and Everett chose to base his interpretation on the Schrödinger equation (instead of path integral or the Wheeler-DeWitt equation, which is only available later in history). Nowadays we understand much better that unless Nature very unexpectedly singles out some time parameter in quantum gravity, a theory based on Schrödinger equation will be wrong [186, 231]. Therefore the Everettian view that a wave function evolving under the Schrödinger equation is all there is [232] is quite likely wrong for the laws of gravity.

One possible rescue is to resort to decoherent histories, which does not rely on the Schrödinger equation [187], to formulate some alternative "Everettian interpretation" [233] in variation from Everett's original one. However, decoherent histories also face issues, as discussed in Section 6.5.2. In addition, present accounts of the Everettian interpretation in decoherent histories [233] are unjustified in imposing the branching structure, due to its time-oriented nature [11]. Additional issues are discussed and debated in [232].

6.5.2 Redundant?

Decoherent histories [187, 234] and Relational quantum mechanics [235, 207] have no trouble accommodating the Standard Model or modern theories of quantum gravity (see [187] for decoherent histories; Chapter 2 of [236] and Section 43.9 of [207] for Relational quantum mechanics) However, they face the danger of redundancy.

Decoherent histories

Gell-Mann and Hartle hold that: "The most general objective of quantum theory is the prediction of the probabilities of individual members of sets of alternative coarse-grained time histories of

[5] See recent works of Tejinder for some preliminary attempts to incorporate the content of the Standard Model into collapse models.

the closed system" [234]. In their decoherent histories interpretation [187, 234], probabilities are assigned to sets of histories obeying the decoherence condition, which ensures the usual probability sum rules under coarse-graining. The histories do not have to refer to experiences, but could, for instance, refer to properties of matter in the early universe before any experiential being came into being. As it currently stands, the decoherent histories interpretation is unable to offer unambiguous predictions. This is due to the lack of a history selection criteria [205]: generically a given history can be embedded in many decoherent sets of histories, and the interpretation offers no rule for selecting among the possibilities. Without a fixed way to extend histories, probabilistic predictions cannot be made unambiguously even for the simple question of what happens next.

The analysis of the previous sections show that as far as empirical predictions are concerned, probability assignment can be restricted to individual experiences instead of histories. There is still the open question to determine the set of possible experiences under a given condition (Section 6.2.2). It is reasonable to hope that a theory of experience can address this question, because if there is going to be a scientific theory of experience at all, as a basic requirement it should tell us what experiences are possible under a given condition.

In contrast, a theory of experience will not reduce the ambiguities associated with decoherent histories. Firstly, as a defining feature, the decoherent histories interpretation encompasses both experiential and non-experiential histories. A theory of experience will not help to reduce the ambiguities associated with non-experiential histories. Secondly, even if one restricts attentions to experiential histories, they will generically refer to multiple experiences. As noted in Section 6.2.3, "joint experience is not experience". It is unclear how a theory of experience can fix ambiguities associated with histories of such non-experiences.

In particular, the ambiguities raised in Section 6.2 appear here as well for experiential histories. Should one assign probability to the histories of multiple experiences (which could be for one being or for multiple beings)? Which set of histories should one use, if one just wants to predict Alice's next experience? One possible way out is to follow the logic of Section 6.2 and assign probabilities only to individual experiences, but not histories. This would suffice to connect theory with experience to meet Task 3 of Section 6.1. In this case, assigning probabilities to histories would be fundamentally redundant. This is not to say that history considerations should be forbidden. Instead, they can still be of practical utility, for as noted by Hartle [237]: "In principle the same prediction could be made from the present data themselves [...] However, it is evidently much easier to start from the event in the past. The reason is that present data contain much information that is irrelevant for this particular future prediction."

Relational Quantum Mechanics

Relational quantum mechanics (RQM) [235, 207] considers "facts" (physical variable taking definite value), which take place whenever two systems interact. Here the systems do not have to be experiential beings, but can be any physical system. Facts are relative to the systems that interact, and are labelled by the interacting systems.

We saw from Section 6.4.1 that with the minimal prescription, quantum theory allows objective empirical predictions: Everyone uses a common formula to predict anyone's experiences. An experience is labelled by the experience itself (e_i, c of (6.4)), and nothing else. In particular, there is no need to refer to systems and their interactions. From this view, RQM is redundant in introducing relative labels where no relative label is needed.

One might argue that there is a gain in the RQM move, because by assigning definite values to variables whenever systems interact, it treats empirical facts and non-empirical facts on equal footing, so that it avoids singling out empirical facts as special. However, in RQM this move only undermined the interpretation itself. For theories of the kind of Section 6.1, RQM holds that "variables actualize at three-dimensional boundaries, with respect to (arbitrary) spacetime partitions" (Section 43.9 of [207]). Here an arbitrary spacetime partition single out two systems, and the interaction at the three-dimensional boundary is supposed to trigger variable actualization at the boundary. Indeed, this does not draw a distinction between empirical facts and non-empirical facts, but it begs the question of whether there are empirical facts at all. Why should one believe that experiences correspond to variables taking definite values on 3D surfaces around 4D regions? Why do experiences not correspond to variables taking definite values, for instance, within 4D regions themselves? Do experiences correspond to variables taking definite values throughout the 3D surfaces? Why not on part of the surfaces? How many 3D surfaces should one consider, for one experience of Alice? ... Without addressing these concerns, RQM's move of treating empirical and non-empirical facts on equal footing obscures how variable actualization relates to experience, and obstructs empirical predictions from the interpretation.

Another side effect of the move is that RQM has to face the preferred basis problem [238, 239]: An interaction does not in general single out a basis for physical variables to take definite values, so with respect to which basis do variables take definite values? Curiously, partially in response to this problem, Adlam and Rovelli now considers a new version of RQM which contains special suppositions for conscious observers [240]: "Therefore in this version of RQM it is now feasible to suppose that the perspective of a conscious observer simply emerges from the collection of the perspectives of all the particles in their brain - roughly speaking, a variable V of a system S will have a definite value v relative to me if variable V has the definite value v relative to most of the particles in my brain (or perhaps just in some particularly relevant section of my brain - we would have to turn to neuroscience to determine how much of the brain should be included)." On the preferred basis problem, Adlam and Rovelli remark that: "RQM need only show that in the limit as one of the systems involved becomes macroscopic, then there is a unique choice of variable which takes definite values in the interaction, in order that macroscopic conscious beings like ourselves can have definite experiences." Decoherence is supposed to pick out the basis, even if "the decoherence process is not perfectly well-defined - there is no exact line between 'decohered' and 'non-decohered'", because "consciousness also does not seem to be perfectly well-defined: to our best current understanding it appears to be some kind of emergent high-level feature of reality, so we are certainly entitled to suppose that consciousness can emerge only when enough decoherence has occurred to single out a well-defined preferred basis." In this new version of RQM to address the preferred basis problem, empirical and non-empirical facts are clearly not on equal footing, so there is no longer any gain in this aspect, in comparison to the minimal prescription of Section 6.4.1. One wonders if there is any gain at all in the RQM move to compensate the redundancy of introducing relative labels discussed above.

6.5.3 Vague?

QBism [221, 223] and neo-Copenhagen [194] interpretations do not assign probabilities to non-empirical histories or facts, so avoids the issue of redundancy. However, they are formulated in vague terms that hinder a direct application to fundamental theories like those in Section 6.1.

QBism

QBism [221, 222, 223] holds that quantum measurement outcomes are just personal experiences for the agents, and quantum probabilities are subjective personal degrees of belief for agents.

QBism has been criticized for being vague and ambiguous (see [241] and references therein). For example, the central concept of agent is loosely characterized as "entities that can take actions freely on parts of the world external to themselves, so that the consequences of their actions matter for them" [217]. This engenders some ambiguities. What counts as "take actions freely"? Does a cell count as an agent? A robot? Alice's brain? Alice's brain in conjunction with Bob's legs? On what basis? One possible way out is to identify individual agents with individual experiential beings. Yet given that joint experience is not experience (Section 6.2.3), this would contradict some previous QBism understandings, e.g., a collection of scientists can act as a single agent (p.1872 of [217]). Ambiguities like this associated with basic notions leave one in wonder what QBism really is saying.

In the context of this paper, another critical issue is whether QBists' focus on the quantum state ψ is misguided. In the understanding of Section 6.4.1, although experiences themselves are subjective, empirical predictions based on (6.4) are objective, in the sense that everyone should use the same formula to predict anyone's experiences. This point is made clear by analyzing the path integral formula (6.4), which refers directly to the fundamental gravity and matter variables. In contrast, QBism focuses on quantum states ψ, which have no fundamental status in the sense explained in Section 6.4.4. If one is free to pick one's favourite ψ to bet on future experiences, certainly some choices will help one gain money, while some will not. It is not false to say that ψ and probabilities derived from it are subjective, but this observation is quite unimportant, if there is formula which always gives the correct empirical predictions to win the bets. Equation (6.4) is supposed to be such a formula, which also makes objective predictions, in apparent tension with QBism.

Neo-Copenhagen interpretation

Brukner's neo-Copenhagen interpretation [194] ascertains the object-subject cut of the traditional Copenhagen interpretation by taking measurement instruments to "lie outside the domain of the theory", and holds that "the quantum state is a representation of knowledge of a (real or hypothetical) observer relative to her experimental capabilities [...] The available experimental precision will in every particular arrangement determine to which objects the observer can meaningfully assign quantum states."

It is not clear how this interpretation applies to theories like those in Section 6.1. What constitutes an "observer" in terms of the fundamental matter and gravity variables? What about an "experiment"? Is there a distinction between experiment and everyday experiences? Does the interpretation accept (6.4) for empirical prediction? If not, what formula does it give? In the brief reference to quantum cosmology in [194], it is mentioned that the observer "is always considered to be external to the universe", since "the 'wave function of the universe' that would include the observer is a problematic concept, as it negates the necessity of the object–subject cut". Does (6.4) with its empirical selection count as treating the observer (or experiential being) external or internal to the universe? It is difficult to infer a definitive answer from [194].

The key notion of "fact" especially deserves a clarification in the neo-Copenhagen interpretation. In the discussion of Wigner's friend settings [195], "facts" which are assigned quantum

probabilities can refer to both "measurement records" and "immediate experiences of observers", and in some statements, "facts" is used interchangeably with "measurement records" and "experiences" [194]: "I will show that any attempt to assume that the measurement records (or "facts" or experiences) that coexist for both Wigner and his friend will run into the problems of the hidden variable program..." However, there is a big difference between record and experience. While joint records is still a record, joint experiences is no longer an experience (Section 6.2). If probabilities are assigned also to measurement records even when no experience refers to them, one is left in wonder what count as records for theories like those in Section 6.1.

Like QBism, the neo-Copenhagen interpretation focuses on quantum states ψ, which have no fundamental status in the sense explained in Section 6.4.4. This is another gap to be filled before one could apply the interpretation to the fundamental theories with specified dynamical laws, if it can be applied to fundamental theories with specified dynamical laws at all.

6.5.4 Superfluous?

The above dangers of being wrong, redundant, or vague are possibly avoided with additional inputs. Yet a much more important issue needs to be tackled first. Are these interpretations worth developing further at all? If the goal is to find a satisfactory physical theory of everything, what difference do the above interpretations make toward this goal?

Alternative theories such as pilot-wave theories and collapse models make different empirical predictions from quantum theory. They also allow for a collective account of experiences in the sense of Section 6.4.2, such that if they succeed, we gain a totally different worldview from quantum theory. These differences they make make them worth developing.

The situation is far less clearer for the other interpretations discussed above. Consider the three tasks given in Section 6.1 for a theory of everything. The interpretations do not tell us the correct dynamical laws for matter and gravity, nor the correct laws for the boundary condition of the universe. The analysis of Section 6.2 gives formula (6.4) to relate theory with experience, without subscribing to any of these interpretations. The outstanding problem is to find a theory of experience which can address the open questions listed in Section 6.2.2. In their current form, these interpretations seem to make no contribution to this task, either. Without contributing anything to the physical theory of everything, these interpretations face the real danger of being superfluous.

6.6 Discussion

Theories of physics are often derived and tested in controlled laboratory settings. This does not mean that the theories only apply to controlled laboratory events. We have learned that physical laws of gravity govern not only Galilei's wood and iron balls, but also birds in free flight. We have learned that physical laws of electromagnetism govern not only Faraday's coils of wire, but also lightnings in thunderstorms.

In this paper, I considered the case that quantum theory applies not only to microscopic phenomena in controlled laboratory settings, but also to everyday experiences. It is assumed that experiences take place probabilistically, i.e., an experience-enabling physical condition c makes

possible a set of experiences $\{e_i\}_i$ to take place with certain probabilities. In path integral theories of everything, the probabilities are given by formula (6.4), reproduced here as

$$p(e_i|c) = N \int_{e_i,c} Dq' \int_{e_i,c} Dq \; e^{\frac{i}{\hbar}(S[q]-S[q'])} \rho(q_b, q'_b).$$

The path integral is doubled to incorporate possibly mixed boundary conditions ρ. The sums include all gravity and matter configurations constrained by the condition c and the experience e_i it enables.[6]

The world picture that emerges from formula (6.4) can be counter-intuitive to some. In accounting for one experience e_i with condition c, we must take it that everywhere else all matter and gravity path integral configurations coexist in superposition. For instance, in the Wigner's friend setting, all other beings must be put in superposition when accounting for one being's experience. This world picture is forced upon us, since by the analysis of Section 6.2.3, the other prescriptions which select path integral configurations differently are not justifiable.

How are different experiences related, if each use of (6.4) only refers to one experience? For the sequential experiences of one being, suppose that previously the experience e_i took place under condition c. This gives rise to another condition c' that enables another set of experiences $\{e'_j\}_j$, whose probabilities are given by (6.4) for $p(e'_j|c')$. The pair (e'_i, c') in turn gives rise to another condition with another set of experiences, etc. Considerations of other beings' experiences are motivated by the content of the experiences of the first being. If in the experience of one being, it perceives the presence of other experiential beings, then it can guess at the conditions d and experiences $\{f_i\}_i$ of the other beings and draw probabilistic predictions for the other beings' experiences using (6.4).

Strictly speaking, no consideration of other beings is needed to account for the experiences of one being. This is implied by (6.4) which holds that all experiences are accounted for by selecting just for that experience. However, in practice we often apply multiple selections in a third-person view (e.g., we apply measurement projections to multiple parties for a Bell experiment) and this often yields valid empirical predictions. How come? The answer is the same as why we often apply non-relativistic analysis which often yield valid empirical predictions. Often a non-relativistic analysis already yields results that meet the desired accuracy. A more consuming relativistic analysis would be an overkill. Likewise, often a third-person view quantum analysis already yields results that meet the desired accuracy. In this case a more consuming first-person analysis based on (6.4) could be saved.

On the other hand, Relativity does not always agree with non-relativistic physics, and it is important to figure out when this happens. Likewise, quantum theory according to (6.4) does not always agree with naive prescriptions, and it is important to figure out when this happens as well. Wigner's friend setting of Section 6.3 is one prominent example. There should be more examples waiting to be explored. For instance, in cosmology, taking (6.4) as the basic formula for empirical predictions will provide a basis to examine how inhomogeneities and anisotropies arise [188], as well as to resolve the Boltzmann brain problem [189]. Specifically, since an experience is characterized completely by the local selection e_i, c of (6.4), the assumption of

[6]The constraint may take the form of a characteristic function, so that gravity and matter configurations compatible with c and e_i are included, while other configurations are excluded. The constraint may also take the form of a function which assigns complex weights to the gravity and matter configurations according to c and e_i, e.g., in implementing "unsharp measurements" induced by the experience. As discussed in Section 6.2.2, the precise form of the constraint hinges on some yet unavailable scientific knowledge about the relationship between physics and experiences.

distinguishability between ordinary and Boltzmann brain experiences based on gravity configurations outside the brains cannot be retained. In another realm, for measurements in QFT, (6.4) enforces a first-person description, which can clarify the issue of superluminal signalling faced by naive measurements prescriptions in a third-person description [190, 191, 192]. Specifically, we must wait for the other agents to send back signals to the agent whose experience is under consideration to distinguish superluminal and non-superluminal cases.

In this paper, I considered (6.4) for arbitrary selections c and e_i. As noted in Section 6.2.2 there are open questions to address if one wants to be more explicit about what e_i and c correspond to actual experiences and conditions. What physical conditions c enable experience? What determines the set of possible experiences $\{e_i\}_i$ given c? What determines the next condition c' given the previous experience e_i and condition c? Are there lawlike features yet to be uncovered about experience? What is the status of "Free Will" in this context? ... Addressing these questions will most require interdisciplinary efforts that push further the boundaries of physics [242, 243, 244, 245, 246, 10]. Intriguingly, a connection appears to have arisen between experience considered here, and Maturana's and Varela's perspectives on cognition [242], stemming from biological studies. Both "experience" and "cognition" stand apart from "observation" or "measurement". The act of observation and measurement inherently assumes the presence of an external entity to be scrutinized, whereas in the realm of experience and cognition, this precondition is absent. Rather than engaging in the quantification or observation of external phenomena, they pertain to a process of self-creation (autopoiesis), perpetuating and nurturing an autonomous living system.

Chapter 7

Indifference boundary condition for the universe

Path integral models for quantum cosmology need boundary condition. An old proposal due to Suen and Young posits that all possible initial boundary configurations should be summed over indifferently. Here we extend the idea in three ways to facilitate the design of empirical tests. Firstly, future boundary conditions are incorporated alongside past conditions, eliminating an inherent past-future asymmetry. Secondly, a mixed-state sum is considered besides the original pure-state sum, leading to a range of candidate boundary conditions for the universe. Lastly, a minimal prescription is employed to exemplify how empirical predictions can be extracted from these boundary conditions. In an application to the de Sitter minisuperspace model, it is shown how the most probable outcome for an observation of the squared scale factor aligns with the saddle points that solve the classical equation of motion. In the picture emerging from the indifference boundary conditions, Big Bang and Big Bounce are not exclusive alternatives. Instead, the path integral includes geometries realizing both possibilities.

7.1 Introduction

If our physical universe is characterized by a path integral, what should be its boundary condition?

This topic has attracted significant research efforts in quantum cosmology [45, 247], including the prominent ideas of the no-boundary proposal and the tunnelling proposal. The no-boundary proposal posits that the universe originates from a non-singular Euclidean geometry at its inception, but faces challenges due to the divergence of the Euclidean gravitational path integral caused by conformal instability, and ambiguities in contour selection when adopting alternative complex contours [109]. The tunnelling proposal posits that the universe is created through a process of quantum tunnelling from a "nothing", but the corresponding path integral is clearly specified only in minisuperspace models with a limited number of degrees of freedom [107, 149], leaving uncertainty on how to apply the proposal to fully general path integrals. Further issues discussed in Section 7.2 surrounding these proposals leads to the prevailing view that the correct boundary condition remains a widely open problem [45, 247].

An alternative proposal due to Suen and Young [114] posits to base quantum cosmology on a path integral over non-singular Lorentzian geometries. The initial boundary condition is such that all possible initial boundary configurations are summed over indifferently. Consequently, this proposal shares with the no-boundary proposal (see Section 7.2.1) the feature that the path integral only includes non-singular spacetimes, thereby achieving singularity avoidance by definition [8]. Additionally, it aligns with the tunneling proposal by utilizing a Lorentzian path integral, thereby circumventing ambiguities associated with the integration contour. In their seminal work, Suen and Young computed the wave function for the universe in the de Sitter minisuperspace model. Interestingly, they found that the wave function approaches zero as the scale factor of the universe diminishes.

Despite the promising characteristics and intriguing results of the original proposal, there has been a surprising lack of subsequent work aimed at testing this proposal for the universe's boundary condition. In this paper, extend the original proposal in three aspects to pave way for deriving empirical predictions which can ultimately be compared against empirical data.

Firstly, we incorporate considerations of the future boundary condition, in addition to the past boundary condition considered originally. This eliminates an apparent past-future asymmetry in the formalism, which seems to lack a fundamental justification.

Secondly, we note that the idea of summing over all possible boundary configurations indifferently can be realized by a mixed-state sum, in addition to the pure-state sum considered originally. By combining both the mixed and pure versions with the past and future boundaries, we derive several distinct boundary conditions, which are collectively referred to as "indifference boundary conditions". Each of these offers a potential candidate for the true boundary conditions of our universe, which enriches the original proposal for a pure-state past condition plus an unspecified future condition.

Among these indifference boundary conditions, the "mixed-mixed condition" realizes a uniform summation over all possible past-future boundary conditions. This can be interpreted as a reasonable prior when approaching the universe's boundary condition as a Bayesian inference problem. Therefore, even if the actual boundary condition of the universe is ultimately different, the indifference boundary condition still provides a credible starting point to discern the true boundary condition via iterative Bayesian updates.

Thirdly, to illustrate a systematic way to extract empirical predictions from the indifference boundary conditions, we utilize the minimal prescription introduce recently [9]. Traditional works of quantum cosmology often stops at computation of the wave function from the past boundary condition, without specifying an explicit prescription for empirical predictions [45, 247]. Employing the minimal prescription exemplifies one way to cover the gap.

In an application to the de Sitter minisuperspace model, it is found that, as expected, the most likely outcome for an observation of the squared scale factor can be inferred from the saddle points that solve the classical equation of motion. Albeit based in a fairly simplified setting, this demonstrates concretely how empirical predictions could be extracted from the supposed boundary condition, and paves the way for further studies in more realistic settings.

In the conventional narrative for cosmology, the Big Bang and the Big Bounce are presented as mutually exclusive alternatives. The indifference boundary conditions suggest that our universe may exist in a superposition of these and other possibilities, all of which are realized by the path integral sum. We take this new vintage point as the most significant conceptual lesson from the study.

The rest of the paper is organized as follows. In Section 7.2, we give a conceptual review of several previous proposals for the boundary condition for the universe. In Section 7.3, we present the indifference boundary conditions in more detail, highlighting how they avoid several shortcomings of alternative proposals. In Section 7.4, we review the minimal prescription for empirical predictions. In Section 7.5, we apply the indifference boundary conditions to the de Sitter minisuperspace model, and show how the most likely outcome for an observation of the squared scale factor can be inferred from the saddle points that solve the classical equation of motion. In Section 7.6, we discuss some outlooks for further research.

7.2 Background

7.2.1 No-boundary proposal

> "Once one allows that singular histories could take part in the path integral, they could occur anywhere and predictability would disappear completely. [...] To implement the idea that the laws of physics hold everywhere, one should take the path integral only over non-singular metrics."
>
> "The path integral for quantum gravity should be taken over all compact Euclidean metrics."
>
> –Stephen Hawking [178]

Hartle and Hawking's no-boundary proposal is based on Euclidean path integrals [44, 177]. Part of the motivation is singularity avoidance [178, 177]. In a Lorentzian metric, $\sqrt{h} = 0$ at some spatial section with spatial metric determinant h implies a singularity. However, the same implication does not hold in the Euclidean case. Consider a 4-sphere of radius R embedded in flat 5-dimensional space, and a 4D slice intersecting the 4-sphere in a 3-sphere of non-zero radius. As the slice shifts towards the boundary, the 3-sphere shrinks to zero volume. Nevertheless, the 4-geometry itself is totally regular there.

This led to the no-boundary proposal that the path integral for quantum gravity should be taken over all compact Euclidean metrics, which are free of singularities. Since the compact Euclidean metrics are also free of boundaries, the boundary condition for the universe is that it has no boundary.

The no-boundary proposal suffers from the conformal instability problem: Since the Euclidean gravitational action is unbounded from below, the path integral does not converge [20]. In practice, one has to consider complex contour inequivalent to the Euclidean one to avoid this divergence issue. As a downside, the above explanation of singularity avoidance in the Euclidean geometric picture becomes inapplicable for a general complex contour. In addition, among the many possible complex contours, the no-boundary proposal fails to single out a unique choice as the boundary condition for the universe [109].

7.2.2 Tunnelling proposal

> "In the sum over histories [...] one has to allow four-geometries with integrable singularities (and finite action), since non-singular compact Lorentzian manifolds do not

exist. [...] Alternatively, one can assume that space-time ceases to be a Lorentzian manifold on scales smaller than $G^{1/2}$."

"$\psi(h, \phi)$ should be obtained by integrating over Lorentzian histories interpolating between a vanishing 3-geometry $\varnothing$ and (h, ϕ) and lying to the past of (h, ϕ) [...]"

"In addition to these path-integral no-boundary proposals, one candidate law of boundary conditions has been formulated directly as a boundary condition in superspace. This is the so-called tunneling boundary condition which requires that ψ should include only outgoing waves at boundaries of superspace. The main weakness of this proposal is that 'outgoing waves' and the 'boundary of superspace' have not been rigorously defined. The Lorentzian path-integral proposal [...] was originally suggested as a path integral version of the tunneling boundary condition, and indeed the two proposals give the same wave function in the simplest minisuperspace model. In the general case, the equivalence of the two proposals is far from being obvious."

–Alexander Vilenkin [142, 248]

Vilenkin's tunnelling proposal originates from the intuitive idea that the universe is spontaneously created at zero or small size by quantum tunnelling from nothing [43]. In the path integral realization of the proposal, the initial size of the universe is set to zero [142]. As specified by the quotation above, the metric changes signature to become singular or to become non-Lorentzian at the initial boundary. In this sense the path integral is only pseudo-Lorentzian, even though it is defined in real time instead of imaginary time.

In the tunneling boundary condition realization of the proposal [148], the boundary of superspace is taken to consist of singular configurations, and is partitioned into two parts. One part includes singular 3-geometries which arise from slicing regular 4-geometries, and is called the "non-singular boundary of superspace". The other part consists of the rest 3-geometries, and is called the "singular boundary of superspace". The tunneling boundary condition requires that at the singular boundary of superspace the wave function includes only outgoing modes.

The tunnelling proposal suffers from several ambiguities. First, the path integral realization of the proposal is not in general equivalent to tunneling boundary condition realization of the proposal [107], which raises the question of which one to choose when they differ. Second, although in simple minisuperspace models with just one scale factor, the path integral realization unambiguously sets the initial scale factor to zero, it is not clear what happens in more general models with multiple scale factors. In models with two scale factors, proposals have been made to set only one scale factor to zero [107, 149]. The situation in more general cases is not clear. Third, as acknowledged in the above quotation, in the tunneling boundary condition realization it is unclear how to define 'outgoing waves' and 'boundary of superspace' precisely beyond the simplest minisuperspace models. For these reasons, it seems appropriate to view the tunnelling proposal as a set of ideas in development instead of a definitive proposal for the boundary condition of the universe.

7.2.3 Indifference proposal

"[The set of path integral configurations] includes all non-singular Lorentzian four-geometries which induce h_{ij} on one of their boundaries, together with all fields ϕ regular on them."

"The boundary condition is taken to be that 'all possible boundaries are included.' "

–Wai-Mo Suen and Kenneth Young [114]

Suen and Young's proposal [114] is built on two key ideas. The first idea is to base quantum cosmology on strictly Lorentzian path integrals where the metric is non-singular and Lorentzian everywhere, including on the boundaries. The second idea is to make no selection on the initial condition, so as to sum over all regular Lorentzian boundary configurations indifferently. To give it a concise name, we will refer to Suen and Young's proposal as the **indifference proposal** in the following.

Because the indifference proposal is defined through Lorentzian path integrals, it does not suffer the ambiguities of the no-boundary proposal based on complex contours. Because the indifference proposal sums over all regular Lorentzian boundaries, it does not suffer the ambiguities of the tunnelling proposal about selecting pseudo-Lorentzian boundary conditions.

On the other hand, the indifference proposal has been much less investigated in comparison to the no-boundary and tunnelling proposals. Beyond the homogeneous and isotropic de Sitter minisuperspace setting studied in the original paper [114], much remains to be investigated about the proposal.

7.2.4 New no-boundary proposal

"Any normalizable wave function (3.1) that satisfies the [Wheeler-DeWitt] constraints is a possible candidate for the quantum state of our Universe from which the probabilities for various types of Lorentzian 4-geometries describing our Universe can be derived."

"A semiclassical no-boundary wave function is defined by a weighted collection of saddle points (extrema) of the action $I[g, \phi]$ on a 4-disk that match (h_{ij}, χ) on its only boundary and are otherwise regular inside."

–Jonathan Halliwell, James Hartle, and Thomas Hertog [249]

Halliwell, Hartle, and Hertog's recent reworking of the no-boundary proposal [249] marks a fundamental departure from the old no-boundary proposal by renouncing any connection to functional integrals. Instead, the wave function of the universe is defined directly as a solution to the Wheeler-DeWitt equation of canonical quantum gravity. At the semiclassical level, the no-boundary wave function is approximated by a weighted sum over saddle points on a 4-disk that match the wave function argument on its only boundary and are otherwise regular inside.

The new no-boundary proposal has the potential of overcoming some shortcomings of the old no-boundary proposal, if one subscribes to canonical quantum gravity. On the other hand, if the goal is to identify suitable boundary condition for the universe in the path integral formalism, then the new no-boundary proposal offers no help.

7.2.5 Other path integral proposals

"Originally, it [the no-boundary proposal] was formulated as a sum over compact and regular geometries. Here the sum is redefined to be over geometries with an approximately zero initial size and approximately Euclidean initial momentum."

–Alice Di Tucci and Jean-Luc Lehners [125]

> "[...] off shell these new definitions all involve sums over universes of various initial sizes, thereby offering the prospect that the physical interpretation of the no-boundary wave function may require further exploration and may end up being even richer than currently assumed."
>
> –Alice Di Tucci, Jean-Luc Lehners, and Laura Sberna [126]

In minisuperspace model path integrals, the original no-boundary proposal is realized by fixing the initial (squared) scale factor to zero. Motivated by considerations of stability issues, Di Tucci, Lehners and Sberna [125, 126] propose to fix the initial momentum instead. According to [125, 126], the no-boundary saddle point can be preserved and the stability issues avoided if the derivative of the squared scale factor is fixed to be Euclidean.

Like the pseudo-Lorentzian path integrals discussed in Section 7.2.2, this alternative path integral implementation of the no-boundary proposal face ambiguities. Firstly, it is yet unclear what the boundary condition should be beyond simple minisuperspace models [250]. Second, already for simple minisuperspace models one needs to consider *ad hoc* integration contours which are neither Euclidean nor Lorentzian to avoid singularities [126]. It remains to be seen how one should select among the abundance of contours in the general case [250]. Finally, a technical concern can be raised about the analysis of the new path integral, which sums the squared scaled factor over the whole real line. It seems reasonable to only allow positive values (or non-negative values if one allows the scale factor to be zero) for the squared scaled factor [4], and it remains to be checked if any conclusion is affected if this change is made.

7.3 Indifference boundary conditions

7.3.1 Motivations

As summarized in Section 7.2.3, the indifference proposal is characterized by two basic ideas. The first is to base quantum cosmology on strictly Lorentzian path integrals where the metric is non-singular and Lorentzian in the bulk and on the boundaries. The second is to sum over all regular Lorentzian boundary configurations indifferently.

As a candidate proposal for the boundary condition of the universe, the indifference proposal is appealing for several reasons. First, there is no ambiguity in the integration contour which is fixed to be Lorentzian. This is in contrast to the original no-boundary proposal, whose complex contour is not uniquely fixed.

Second, although the original article [114] focused on minisuperspace models, the two ideas summarized above can be applied directly in fully general gravitational path integrals, such as the Lorentzian simplicial quantum gravity [2], which accommodate inhomogeneity and anisotropy beyond minisuperspace models [72], and readily implements the idea of summing over non-singular spacetime geometries [8]. This is in contrast to the tunnelling proposal, whose formulation beyond minisuperspace is unclear.

Third, summing over all allowed boundary configurations avoids the need to pick a particular configuration. The no-boundary and tunnelling path integrals invokes a vanishing boundary size for the path integral. However, since with a vanishing size the metric changes signature, this

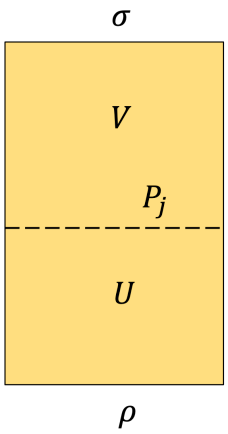

Figure 7.1: A set of outcomes represented by projectors $\{P_j\}$ inside a universe with past boundary condition ρ, future boundary condition σ, and unitary evolution U, V.

option is unavailable in a strictly Lorentzian path integral. Yet any other choice for the boundary size, such as fixing the boundary size of the universe to be 42, appears *ad hoc*. Summing over all boundary configurations has the merit of avoiding such a choice.

In the remainder of this paper, we focus on the indifference proposal. We will engage in some groundwork, aimed in particular at providing a prescription for empirical predictions, which can be ultimately applied to test the proposal against data.

7.3.2 The need for future boundary condition

Quantum theory allows the imposition of both past and future boundary conditions [251, 252, 253]. For example, consider the situation illustrated in Figure 7.1. In the interior of the universe a set of alternative outcomes occur with quantum probabilities

$$p_j = \mathcal{N}\,\mathrm{Tr}\Big[\sigma V P_j U \rho\, U^\dagger P_j V^\dagger\Big], \tag{7.1}$$

where ρ, σ are the density operators for the past and future boundary conditions, U, V are the unitary evolution operators, $\{P_j\}$ are the projection operators for the outcomes, and $\mathcal{N}$ is a normalization constant.

However, textbook applications of quantum mechanics often ignore future boundary conditions and use

$$p_j = \mathcal{N}\,\mathrm{Tr}\Big[P_j U \rho\, U^\dagger P_j\Big] \tag{7.2}$$

for the probabilities. If applied to quantum cosmology, (7.2) would assume from the outset that there is a past boundary condition, but not a future boundary condition for the universe. This presumes a special form of time asymmetry, which may not hold in Nature. One possibility to reconcile (7.1) and (7.2) is to take the future condition

$$\sigma \propto I \tag{7.3}$$

to be proportional to the identity [253]. Then (7.1) reduces to (7.2) using cyclicity of the trace. Since (7.1) includes (7.2) as a special case with $\sigma \propto I$, we adopt the more general formula of (7.1) with both past and future boundary conditions in the following study. In this context, to discover the boundary condition for the universe means to discover both ρ and σ.

7.3.3 Versions of indifference boundary condition

In a framework that allows for both past and future conditions, as well as both pure and mixed boundary conditions, there are multiple ways to implement the indifference sum over boundary configurations.

If we take the past or future boundary condition to be pure, then the idea of including all possible boundaries indifferently can be realized by $|\phi\rangle = \sum_k |k\rangle$, which sums over all vectors of an orthonormal basis indifferently. The corresponding (unnormalized) density operator is

$$\phi = |\phi\rangle\langle\phi| = \sum_{k,l} |k\rangle\langle l| \,. \tag{7.4}$$

If we take the past or future boundary condition to be mixed, then the idea of including all possible boundaries indifferently can be realized by the (unnormalized) density operator

$$\omega = \sum_{k} |k\rangle\langle k| \,, \tag{7.5}$$

which sums over all states of an orthonormal basis incoherently and indifferently. In quantum cosmology, this boundary condition has been discussed by Gell-Mann and Hartle as the "condition of indifference" [253].

In their original study for the de Sitter minisuperspace model [114], Suen and Young considered path integrals of the form

$$\psi(a_0) = \int \mathrm{d}a_i \int DN \int_{a_i}^{a_0} Da \; e^{iS[a,N]} \tag{7.6}$$

for the wave function evaluated at scale factor a_0. Here the functional integral $\int_{a_i}^{a_0} Da$ sums over scale factor histories beginning at a_i and ending at a_0. The path integral for the lapse $\int DN$ sums over different proper times for the histories. The integral $\int \mathrm{d}a_i$ sums all non-singular beginning scale factors a_i over $(0, \infty)$. This last step implements the pure state sum (7.4) for the past boundary.

When both future boundary and mixed conditions are incorporated, the path integral should be updated into a "double path integral" with the partition function (see e.g., equation (4.8) of [187])

$$Z[\rho, \sigma] = \mathcal{N} \int Dx' \int Dx \; \sigma(x_f, x'_f) e^{i(S[x] - S[x'])} \rho(x_p, x'_p). \tag{7.7}$$

This path integral for general physical variables x may be viewed as the path integral version of (7.1) without the projection. As illustrated in Figure 7.2, histories between the past and future boundaries are integrated twice, once in $\int Dx$ weighted by $e^{iS[x]}$, and once $\int Dx'$ weighted by $e^{-iS[x']}$. In $\int Dx$ and $\int Dx'$, all past and future configurations are summed over with complex weights assigned by the past and future boundary conditions ρ and σ.

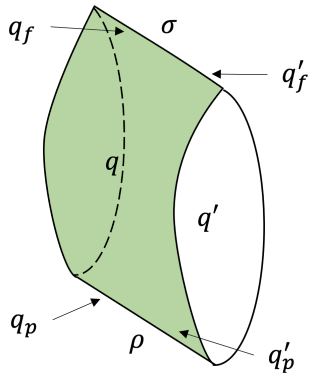

Figure 7.2: In a double path integral, the path q with past boundary x_p and future boundary x_f is integrated on the front sheet, while the path q' with past boundary x'_p and future boundary x'_f is integrated on the back sheet. The two sheets interact through the past and future boundary conditions $\rho(x_p, x'_p), \sigma(x_f, x'_f)$.

As functional versions of (7.4) and (7.5), we have

$$\phi(x_b, x'_b) = 1, \quad \forall x_b, x'_b, \tag{7.8}$$
$$\omega(x_b, x'_b) = \delta(x_b, x'_b), \tag{7.9}$$

where $x_b = x|_B$ and $x'_b = x'|_B$ are values of x and x' at the past or future boundary B. Depending on the choice of pure or mixed sum, we have the following versions of the indifference proposal for the boundary condition of the universe:

$$\textbf{Pure-pure condition: } \rho = \phi,\ \sigma = \phi, \tag{7.10}$$
$$\textbf{Mixed-mixed condition: } \rho = \omega,\ \sigma = \omega, \tag{7.11}$$
$$\textbf{Pure-mixed condition: } \rho = \phi,\ \sigma = \omega, \tag{7.12}$$
$$\textbf{Mixed-pure condition: } \rho = \omega,\ \sigma = \phi. \tag{7.13}$$

Although we focus on the above four cases below, it is in principle possible to choose some other condition for one of ρ and σ to implement the indifferent proposal only on the past of future boundary. Moreover, if the boundary of the universe has timelike components [206], it is possible to introduce pure or mixed sum over timelike configurations on these components. In general, the boundary of the universe may have one or multiple components, with or without designated causal signatures, and there would be indifference boundary conditions that sum over pure or mixed configurations on these components, with or without designated causal signatures accordingly.

7.3.4 Inferential perspective

The above indifference boundary conditions can be viewed as candidates for the true boundary condition for the universe. However, there is an alternative **inferential perspective**, which motivates the consideration of a particular indifference boundary condition, even if the true boundary condition for the universe is something else.

The question of the boundary condition for the universe can be approached from an inferential perspective, much akin to examining the properties of a coin. In the absence of any prior

knowledge, one can postulate a prior probability distribution, which in this case would be a uniform distribution analogous to considering a fair coin where the chances of landing on either side are equal. In case the coin is not actually fair, this initial distribution is subject to updates as new information is accumulated in the manner of Bayesian inference.

In a similar vein, the mixed-mixed condition (7.11) may be viewed as a uniform distribution over all possible boundary conditions for the universe. This distribution serves as an appropriate prior to commence the inference of the unknown boundary condition of the universe. Predictions can be drawn from this prior to compared with new information and data. If other boundary conditions turns out to fit better, the prior can be updated in the manner of Bayesian inference. In this perspective, the mixed-mixed indifference boundary condition serves as a reasonable starting point for systematic inference of the universe's boundary condition.

7.4 Empirical predictions

7.4.1 The need for interior condition

> "In practice, one is normally interested in the probability, not of the entire 4-metric, but of a more restricted set of observables. Such a probability can be derived from the basic probability (2.1) by integrating over the unobserved quantities. In cosmology, one is concerned with observables, not at infinity, but in some finite region in the interior of the 4-geometry."
>
> –Stephen Hawking [177]

All of (7.10) to (7.13) constitute candidates for the boundary condition of the universe. In addition, the mixed-mixed condition (7.11) can be viewed as a prior for inferring the boundary condition of the universe. To test these conditions against data, we must give a prescription to extract empirical predictions from them.

In quantum cosmology, the whole universe is treated as a quantum system, and observations are made in the interior of this quantum system. This is in contrast to particle physics, where observations associated with the S-matrix are located on the outer-boundary of the quantum system under study. Therefore in quantum cosmological path integrals such as (7.7), **interior conditions** need to be supplied in addition to boundary conditions to derive probabilistic predictions.

By interior condition, we mean to introduce an additional complex-valued **weight function** w into the path integral to obtain

$$Z[\rho,\sigma,w] = \mathcal{N}\int Dx' \int Dx\ \sigma(x_f,x'_f)e^{i(S[x]-S[x'])}w[q,q']\rho(x_p,x'_p). \tag{7.14}$$

The form of w depends on the specific event under consideration. For instance, to imitate (7.1) for a sharp measurement, we should exclude paths incompatible with a measurement outcome from the path integral when deriving probability for that outcome. In this case w is a characteristic function that assigns 0 to paths to be excluded, and 1 to paths to be included. More generally, unsharp measurements will invoke w with a continuous range. For situations involving multiple events/measurements, it is not *a priori* clear how w should be chosen. Below we will review two general schemes to systematically incorporate interior conditions.

7.4.2 Decoherent histories

> "The most general objective of quantum theory is the prediction of the probabilities of individual members of sets of alternative coarse-grained time histories of the closed system."
>
> –Murray Gell-Mann and James Hartle [234]

One example for a framework that incorporates interior conditions is the decoherent histories framework. In this formalism, probabilities are assigned to alternative histories obeying the so-called decoherence condition, which ensures the usual probability sum rules under coarse-graining [254, 253, 187] (see also [252, 255]).

To illustrate the idea, a sequence of n quantum measurements, where a particular history of outcomes $\alpha = \alpha_1, \cdots, \alpha_n$ occurs with probability

$$p(\alpha) = \mathcal{N} \operatorname{Tr}[\sigma P_{\alpha_n}(t_n) \cdots P_{\alpha_1}(t_1) \rho P_{\alpha_1}(t_1) \cdots P_{\alpha_n}(t_n)]. \tag{7.15}$$

This is just a multiple measurement generalization of (7.1), with P as the projection operators for the outcomes. For simplicity of notation we adopted the Heisenberg picture, where an operator $O(t)$ at time t is related to the Schrödinger picture operator O by $O(t) = U^\dagger(t, t_0) O U(t, t_0)$ for the unitary evolution $U(t, t_0)$ from reference time t_0 to t. To simplify the notation further one defines the operators

$$C_\alpha = P_{\alpha_n}(t_n) \cdots P_{\alpha_1}(t_1). \tag{7.16}$$

Then we have $p(\alpha) = D(\alpha, \alpha)$ in terms of the decoherence functional

$$D(\alpha, \beta) = \mathcal{N} \operatorname{Tr}\left[\sigma C_\alpha \rho C_\beta^\dagger\right]. \tag{7.17}$$

For path integrals, the decoherence functional can be defined as [187]

$$D(\alpha, \alpha') = \mathcal{N} \int_{\alpha'} Dx' \int_{\alpha} Dx \; \sigma(x_f, x'_f) e^{i(S[x] - S[x'])} \rho(x_p, x'_p). \tag{7.18}$$

Here the set of paths are classified into subsets labelled by α. The first integral sums over paths in α', and the second integral sums over paths compatible with α.

In the decoherent histories framework, one focuses on sets of histories $\{\alpha\}$ that obey the additional decoherence condition

$$D(\alpha, \beta) = 0, \quad \alpha \neq \beta, \tag{7.19}$$

which ensures the additivity property for probabilities under coarse-graining of histories. The additivity property for a set of probabilities $\{p(\alpha)\}$ says that $p(\alpha) + p(\alpha') = p(\alpha \vee \alpha')$ for $\alpha \neq \alpha'$. One can check that (7.19) is sufficient to imply that the above property holds with $p(\alpha \vee \alpha') = \mathcal{N} \operatorname{Tr}\left[\sigma C_{\alpha \vee \alpha'} \rho C^\dagger_{\alpha \vee \alpha'})\right]$ for $C_{\alpha \vee \alpha'} = C_\alpha + C_{\alpha'}$.

Gell-Mann and Hartle's decoherent histories framework originates from considerations of quantum cosmology [254], and there has been attempts to apply the formalism to extract probabilistic predictions for quantum cosmology (e.g. [256, 257, 258, 259]). However, the formalism suffers some long-standing issues [205]. In particular, a given event can belong to many different histories, and a given history can belong to many different decoherent sets of histories, but the formalism does not offer a rule to select the set of histories to arrive at definite probabilistic predictions for the event. This motivates us to consider an alternative formalism next.

7.4.3 Minimal prescription

The minimal prescription [9] provides an alternative to incorporate interior conditions in path integrals to arrive at empirical predictions. Unlike the decoherent histories formalism, whose histories do not have to refer to experiences, the minimal prescription focuses on the experiences of experiential beings, i.e., beings with experience, such as human beings.

In a world where experiences unfold deterministically, making empirical predictions amounts to determining which experience will occur under a given condition. More generally, if more than one experience is possible to occur under a given condition, we need to consider conditional probabilities

$$p(e_i|c), \tag{7.20}$$

where c stands for the condition, and $\{e_i\}_i$ stands for the set of possible experiences. When a unique experience is possible, the set is a singleton as a special case.

In a comprehensive enough physical theory, $p(e_i|c)$ should be derived as empirical predictions of the theory. The **minimal prescription** for path integrals assumes

$$p(e_i|c) = \mathcal{N} \int Dx' \int Dx\ \sigma(x_f, x'_f) e^{i(S[x]-S[x'])} w_{e_i,c}[q, q'] \rho(x_p, x'_p) \tag{7.21}$$

for the form of $p(e_i|c)$. This is just (7.14), with the weight function determined and labelled by the condition c as well as experiences e_i.

The minimal prescription is "minimal", because the path integral for $p(e_i|c)$ is weighted only by $w_{e_i,c}$ and nothing else. An experiential being such as a human being can have many experiences before and after e_i. These experiences have other weight functions. According to the minimal prescription, none of them show up in the path integral for $p(e_i|c)$. If a previous experience influences the probability for e_i, it can only be through c. For example, the condition c could refer to the brain physical configuration of a human being, specifying the memory of previous experiences. Such memories can influence e_i, in the sense that a different memory with c' could change the probability, i.e.,

$$p(e_i|c) \neq p(e_i|c'). \tag{7.22}$$

Conversely, if a previous experience is completely forgotten and leaves no trace in c, then it cannot influence the future experiences, i.e.,

$$p(e_i|c, f) = p(e_i|c, f') \tag{7.23}$$

for previous experiences $f \neq f'$. This minimal aspect contrasts sharply with the decoherent histories whose weight functions depend on previous and later events, and thus avoids the issue of the lack of selection criterion for embedding events in histories.

An outstanding task for the minimal prescription program is to determine the form of $w_{e_i,c}$ for all experiences and condition. This task is hard because it touches on the question of translating between physical configurations and conscious experience. Our current scientific understanding is unable to specify the precise set of experiences $\{e_i\}_i$ possible under a given physical condition c. For the present study of quantum cosmology, we shall assume a very high level understanding of w. For a sharp measurement, $w_{e_i,c} \to \{0, 1\}$ is a characteristic function, assigning 1 to all

matter and spacetime configurations where the physical configurations for the condition c and experience e_i of seeing the sharp measurement outcome are fulfilled somewhere. Otherwise $w_{e_i,c} = 0$. For an unsharp measurement, $w \in \mathbb{C}$ take more general values, but peak around the experience of seeing the unsharp measurement outcome. Although the details are missing for what e_i and c concretely are, this structural-level understanding suffices for the study below, which establishes a connection between saddle point(s) and leading order empirical prediction in a toy models universe.

7.5 Example: de Sitter minisuperspace

7.5.1 De Sitter minisuperspace

In this section we apply the indifference boundary condition in the de Sitter minisuperspace model to derive empirical predictions using the minimal prescription. We will see that the most likely outcome for an observation of the squared scale factor can be inferred from the saddle points that solve the classical equation of motion. We will go through the steps meticulously even though this makes the derivation quite lengthy for some easily expected results, since it is worth illustrating how the results emerges systematically from the prescription rather than from loose intuitions.

The de Sitter minisuperspace model is described by the spatially closed minisuperspace metric [139, 105, 60]

$$\mathrm{d}s^2 = -\frac{N^2}{q(t)}\mathrm{d}t^2 + q(t)\left(\frac{1}{1-r^2}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2\right)\right) \tag{7.24}$$

with squared scale factor $q(t)$, lapse N, and positive spatial curvature k.

The action for the cosmological constant dominated universe without matter is given by the Einstein-Hilbert action plus the Gibbons-Hawking-York boundary term [140, 141]: $S = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}(R-2\Lambda) + \int_B \mathrm{d}^3y\sqrt{h}K$, where $\Lambda > 0$. Plugging (7.24) in this action yields

$$S[q,N] = 2\pi^2 \int_0^1 \mathrm{d}t\, N\left(-\frac{3}{4N^2}\left(\frac{\mathrm{d}q}{\mathrm{d}t}\right)^2 - \Lambda q + 3k\right). \tag{7.25}$$

Here we have taken the $dN/dt = 0$ gauge for N [139, 105]. In term of the physical proper time $\tau = Nt$,

$$S[q,N] = 2\pi^2 \int_0^N \mathrm{d}\tau \left(-\frac{3}{4}\dot{q}^2 - \Lambda q + 3k\right), \tag{7.26}$$

where dot denotes derivative with respect to τ. This form of the action allows us to see straightforwardly the physical meaning of N as the proper time separation between the past and future boundaries.

At the classical level, the model is solved by the quadratic function

$$q(\tau) = \frac{1}{3}\Lambda\tau^2 + c_2\tau + c_1, \tag{7.27}$$

where c_1, c_2 are constants to be fixed by boundary conditions. For a positive cosmological constant Λ, the universe expands with positive acceleration. If we start at a large q and trace backwards in time, the universe shrinks to a small size and then bounces. At the quantum level, we adopt the minimal prescription to derive empirical predictions next.

7.5.2 Weight function

Imagine an experience of seeing the squared scale factor to be peaked around q_2, conditioning on seeing it peaked around q_1 a proper time $\Delta\tau$ earlier. We want to derive probabilities for such experiences of different q_2 in the minisuperspace model.

In the language of Section 7.4.3, we are considering

$$c = q_1, \Delta\tau, \quad \{e_i\} = \{q_2\}, \tag{7.28}$$

where $q_1, \Delta\tau$ are fixed, while different experiences are parametrized by different q_2 values. To describe q peaking around a certain value q_0 at time τ_0, we introduce the normalized Gaussian distribution

$$\chi[q; \tau_0, q_0, u] = \frac{\sqrt{\pi}}{u} e^{-\frac{(2\pi^2)}{2}\left(\frac{q(\tau_0)-q_0}{u}\right)^2}. \tag{7.29}$$

This suppresses a path $q(\tau)$ if at τ_0 the value of q is away from q_0, and enhances it if q is close to q_0. The weight functions are given by

$$w_{c,q_2}[q, q'] = v_{c,q_2}[q] v^*_{c,q_2}[q'], \tag{7.30}$$

$$v_{c,q_2}[q] = \int_0^{N-\Delta\tau} \chi[q; \tau_1, q_1, u_1]\ \chi[q; \tau_1 + \Delta\tau, q_2, u_2]\ \mathrm{d}\tau_1, \tag{7.31}$$

Here w factorizes into the components v for paths q, q' on the first and second sheets of the double path integral. The two χ check the requirements at τ_1 and $\Delta\tau$ later, and suppresses or enhances q accordingly. Since no observation is made of time, the requirements can be fulfilled at any $\tau_1, \tau_1 + \Delta\tau \in (0, N)$. Therefore τ_1 is integrated over, and the upper limit $N - \Delta\tau$ ensures that $\tau_1 + \Delta\tau < N$. For simplicity, in the following we suppress the subscript c, q_2 when there is no ambiguity.

The probabilities are derived by plugging (7.30) in (7.21). For further analysis, it is convenient to separate out the boundary value integrals of (7.21) to re-express it as

$$p(q_2|q_1, \Delta\tau) = \mathcal{N} \int Dx'_f Dx'_p Dx_f Dx_p\ \tau(x_f, x'_f)\rho(x_p, x'_p)\ Z[x_p, x_f, x'_p, x'_f, w], \tag{7.32}$$

where we defined

$$Z[x_p, x_f, x'_p, x'_f, w] = \int_{x'_p}^{x'_f} Dx' \int_{x_p}^{x_f} Dx\ e^{i(S[x]-S[x'])} w[q, q']. \tag{7.33}$$

Since w of (7.30) factorizes, $Z[x_p, x_f, x'_p, x'_f, w]$ in turn factorizes as

$$Z[x_p, x_f, x'_p, x'_f, w] = Z[x_p, x_f, v] Z^*[x'_p, x'_f, v], \tag{7.34}$$

$$Z[x_p, x_f, v] = \int_{x_p}^{x_f} Dx\ e^{iS[x]} v[q]. \tag{7.35}$$

Plugging in (7.31) for v, we can express $Z[x_p, x_f, v]$ in terms of the new "action" R as

$$Z[q_p, q_f, v] = \frac{\pi}{u_1 u_2} \int DN \int_0^{N-\Delta\tau} \mathrm{d}\tau_1 \int_{q_p}^{q_f} Dq\ e^{iR[q,N,\tau_1]}, \tag{7.36}$$

$$R[q, N, \tau_1] = 2\pi^2 \left(\int_0^N \mathrm{d}\tau \left(-\frac{3}{4}\dot{q}^2 - \Lambda q + 3k \right) + \frac{i}{2} \left(\frac{q(\tau_1) - q_1}{u_1} \right)^2 + \frac{i}{2} \left(\frac{q(\tau_1 + \Delta\tau) - q_2}{u_2} \right)^2 \right). \tag{7.37}$$

Equation (7.36) assumes a functional integral form for variables N, τ_1, q. We can perform a saddle point analysis to identify the leading contribution to the integral.

7.5.3 Saddle points

Variation of $Z[x_p, x_f, v]$ with respect to q yields

$$\left(\frac{3}{2}\ddot{q} - \Lambda\right)\delta q + \frac{3}{2}\dot{q}\delta q|_{\tau=0}^{\tau=N} + i\left(\frac{q-q_1}{u_1^2}\delta q|_{\tau_1} + \frac{q-q_2}{u_2^2}\delta q|_{\tau_1+\Delta\tau}\right) = 0 \tag{7.38}$$

For the Dirichlet boundary condition $q(0) = q_p, q(N) = q_f$ of (7.36), this is solved by

$$q(\tau) = \frac{\Lambda}{3}\tau^2 + \left(-\frac{\Lambda}{3}N^2 + q_f - q_p\right)\frac{\tau}{N} + q_p, \tag{7.39}$$

$$q(\tau_1) = q_1, \quad q(\tau_1 + \Delta\tau) = q_2. \tag{7.40}$$

Variation with respect to N yields [139, 105, 60]

$$\frac{3}{4}\dot{q}^2 + 3k - \Lambda q = 0. \tag{7.41}$$

Given (7.39), this is solved by

$$N_\pm = \frac{3}{\Lambda}\left((\frac{\Lambda}{3}q_p - k)^{1/2} \pm (\frac{\Lambda}{3}q_f - k)^{1/2}\right). \tag{7.42}$$

Negative (7.42) also solves the equation. However, this leads to the same metric in (7.24), so we only consider (7.42). Plugging $N_\pm$ in (7.39) yields

$$q_\pm(\tau) = \frac{\Lambda\tau^2}{3} \mp \frac{2}{3}\sqrt{3\Lambda q_p - 9k}\,\tau + q_p. \tag{7.43}$$

One can check that when $\tau = N_\pm$, $q_\pm = q_f$.

When $q_p, q_f > 3k/\Lambda$, $N_\pm$ are real. The quadratic functions (7.43) are portions of parabolas with longer and shorter proper times (Figure 7.3). This reflects the fact that the U-shaped parabola can meet the same boundary condition twice before and after the turning point. Therefore in (7.43), q_- can actually be viewed as a monotonic portion of q_+: When $q_p > q_f$, q_- covers the decreasing portion toward the beginning of q_+; When $q_p < q_f$, q_- covers the increasing portion towards the ending of q_+; When $q_p = q_f$, q_- has zero extension in proper time since $N_- = 0$. Note also from (7.43) that for the longer path, the minimum of q is always at $3k/\Lambda$.

When either $q_p < 3k/\Lambda$ or $q_f < 3k/\Lambda$, $N_\pm$ becomes imaginary, leading to complex saddle points. These classically forbidden paths are meaningful in characterizing quantum tunnelling events where q reaches classically forbidden small values. However, they cannot satisfy (7.40) for $q_1, q_2 \in \mathbb{R}$. Therefore we focus on the case of $q_p, q_f > 3k/\Lambda$ below.

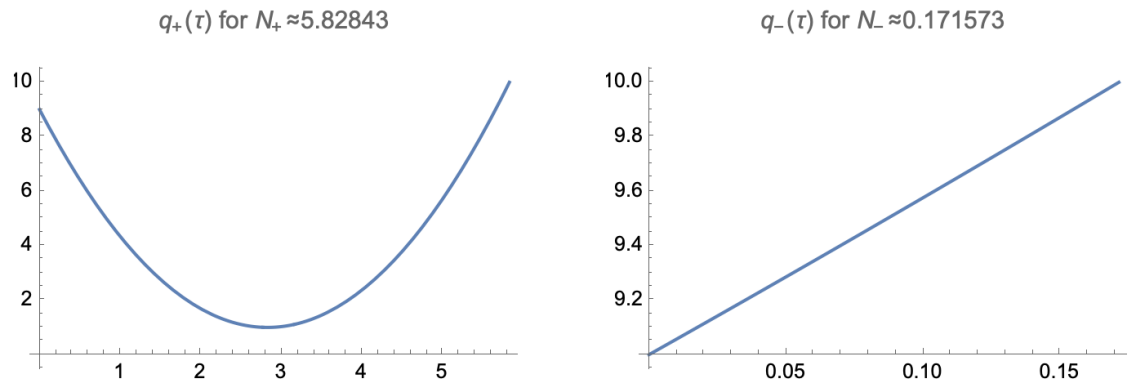

Figure 7.3: $q_\pm(\tau)$ for $k = 1, \Lambda = 3, q_p = 9, q_f = 10$

7.5.4 Most likely experience

The weight functions of Section 7.5.2 yield probabilities

$$p(q_2|q_1, \Delta\tau) \tag{7.44}$$

through (7.32) for the experience of observing the squared scale factor q_2 when at proper time $\Delta\tau$ earlier the squared scale factor was observed to be q_1. We apply the saddle point analysis of Section 7.5.3 to identify the value of q_2 with the largest probability, i.e.,

$$q_2^* = \underset{q_2}{\operatorname{argmax}}\, p(q_2|q_1, \Delta\tau). \tag{7.45}$$

For given $q_1, \Delta\tau$, this is the value(s) of q_2 obeying (7.40), because it lies on saddle points which dominate the path integral [60, 4].

As noted above, q_- can be viewed as a monotonic portion of q_+. Therefore for given $q_1, q_2, \Delta\tau$ values, whether some saddle point p_- obeys (7.40) can be inferred by looking at the corresponding portions of q_+. Hence without loss of generality we will focus on q_+ in identifying values of q_2 meeting (7.40).

Consider the first condition of (7.40), which allows us to identify saddle point paths crossing q_1. Solving $q_+(\tau) = q_1$ for τ_1, we obtain two candidate solutions

$$\tau_a = \frac{\sqrt{3}\left(\sqrt{\Lambda q_p - 3k} - \sqrt{\Lambda q_1 - 3k}\right)}{\Lambda}, \tag{7.46}$$

$$\tau_b = \frac{\sqrt{3}\left(\sqrt{\Lambda q_1 - 3k} + \sqrt{\Lambda q_p - 3k}\right)}{\Lambda}, \tag{7.47}$$

which may or may not belong to the domain $[0, N]$. The U-shaped quadratic function $q_+(\tau)$ has maximum at its side at q_p or q_f, and minimum at $3k/\Lambda$, as noted above. Therefore when $q_p \le q_f$,

$$\tau_1 = \begin{cases} N/A, & q_f < q_1, \\ \tau_b, & q_p < q_1 \le q_f, \\ \tau_a \text{ or } \tau_b, & 3k/\Lambda \le q_1 \le q_p, \\ N/A, & q_1 < 3k/\Lambda. \end{cases} \tag{7.48}$$

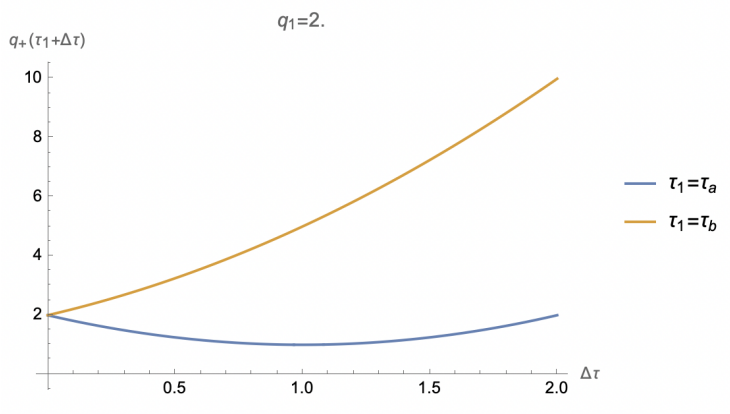

Figure 7.4: $q_+(\tau_1 + \Delta\tau)$ for $k = 1, \Lambda = 3, q_1 = 2$

When $q_p > q_f$,

$$\tau_1 = \begin{cases} N/A, & q_p < q_1, \\ \tau_a, & q_f < q_1 \leq q_p, \\ \tau_a \text{ or } \tau_b, & 3k/\Lambda \leq q_1 \leq q_f, \\ N/A, & q_1 < 3k/\Lambda. \end{cases} \tag{7.49}$$

For q_1 values leading to N/A, there will not be a saddle point crossing q_1 to meet (7.40). In other cases, the first condition of (7.40) is met at some $\tau_1 = \tau_a$ or τ_b. To check the second requirement of (7.40) at proper time $\Delta\tau$ later, consider

$$q_+(\tau_1 + \Delta\tau) = \begin{cases} q_1 + \frac{1}{3}\Delta\tau\left(\Delta\tau\Lambda - 2\sqrt{3\Lambda q_1 - 9k}\right), & \text{if } \tau_1 = \tau_a, \\ q_1 + \frac{1}{3}\Delta\tau\left(\Delta\tau\Lambda + 2\sqrt{3\Lambda q_1 - 9k}\right), & \text{if } \tau_1 = \tau_b. \end{cases} \tag{7.50}$$

This single out the values of q_2 meeting (7.40). A sample of these functions is plotted in Figure 7.4. Since τ_a is prior to the turning point, the value decreases first before it increases, and since τ_b is after the turning point, the value directly increases.

Now let us put all the pieces together and summarize the physical meaning of the above considerations. The probabilities $p(q_2|q_1, \Delta\tau)$ are derived from a strictly Lorentzian path integral (7.32) over non-singular geometries with $q(\tau) > 0$ [4], as advocated in Suen and Young's original proposal [114]. Under any version from (7.10) to (7.13) of the indifference boundary conditions, the path integral sum over all boundary values $q_p, q_f > 0$ with non-vanishing support. Hence the saddle points of $Z[q_p, q_f, v]$ of (7.36) for any choice of $q_p, q_f > 0$ are relevant. From (7.48) and (7.49), we see that no saddle point covers q_1 when $q_1 < 3k/\Lambda$. If the previous observation yielded such a highly unlikely outcome for an extremely small universe, the saddle-point analysis would not be able to inform us the most likely outcome for the next measurement. However, as long as

$$q_1 \geq 3k/\Lambda, \tag{7.51}$$

there are always large enough q_p, q_f such that q_1 belongs to some saddle point according to (7.40). In this case, the most likely outcome(s) q_2^* for the next observation is determined by (7.50), where whether τ_1 equals τ_a, τ_b or both is determined by (7.48) and (7.49).

This most likely experience for the value q_2 at $\tau + \Delta\tau$ agrees with the value one would infer from solving the classical equation of motion, given the value of q_1 at τ. Note that (7.50) is independent of q_p and q_f, so q_2 that peaks $p(q_2|q_1, \Delta\tau)$ in fact lies on infinitely many saddle points with different boundary values q_p and q_f.

Note also that the same result holds to all versions of indifference boundary conditions. As far as the most likely observation for q_2 goes, these boundary conditions yield indistinguishable empirical predictions in the de Sitter minisuperspace model.

The whole setup may also be viewed as a measurement to determine the value of the cosmological constant Λ. If we observed the squared scale factor to be q_1 first, and at proper time $\Delta\tau$ the data for the observational data for the squared scale factor is peaked around the value q_2. Then from these three numbers we can solve for Λ using (7.50). This view also helps us understand why in the present model it is difficult to distinguish versions of indifference boundary conditions from each other: these candidate boundary conditions share the same Λ.

7.6 Discussion

Path integrals for quantum cosmology and quantum gravity need boundary conditions. The much-studied no-boundary and tunnelling boundary conditions are beset with persistent issues reviewed in Section 7.2. For instance, the no-boundary proposal faces ambiguities in the choice of integration contour, while the tunnelling proposal remains unspecified beyond simple minisuperspace settings.

In the alternative proposal put forth by Suen and Young [114], all initial boundary configurations are summed over indiscriminately in a path integral over non-singular Lorentzian geometries. The present work builds on this original proposal and extends it in three aspects. Firstly, the concept of summing over all configurations is now applied to the future boundary, in addition to the past boundary. Secondly, an alternative mixed-state sum over all configurations is introduced, in addition to the original pure-state sum. Combining the mixed and pure versions with the past and future boundaries leads to several different boundary conditions as enumerated in (7.10) - (7.13). These are collectively referred to as "indifference boundary conditions". Thirdly, the minimal prescription [9] is employed to illustrate a method for deriving empirical predictions from the indifference boundary conditions. This extends beyond merely computation of the wave function from the past boundary condition [114], which left the derivation of empirical predictions as an open question. Based on the minimal prescription, the indifference boundary conditions are applied to de Sitter minisuperspace model for pure gravity. It is found that, as expected, the most likely experience for observations of the squared scale factor can be inferred using the classical solution.

These results illustrate how empirical predictions can be obtained from cosmological boundary condition proposals including the indifference boundary conditions. In addition, they show how the indifference boundary conditions passes the test for producing an expected empirical prediction in de Sitter minisuperspace model. However, the results from the fairly simple minisuperspace model are not enough to distinguish versions of indifference boundary conditions from each other. Future research should naturally pivot towards exploring further tests for the indifference boundary conditions that preferably distinguish them among themselves, and from other boundary condition proposals.

As an example, one immediate topic of interest concerns the arrow of time. Traditional approaches, such as the past hypothesis [260, 261], attempt to explain the thermodynamic arrow of time by positing a low-entropy initial boundary condition. However, this avenue certainly becomes inaccessible for the mixed-mixed indifference boundary condition: Not only does the initial boundary condition have maximum instead of low entropy, but the past and future boundary conditions are also identical, eliminating any inherent asymmetry at the level of the boundary condition. If we assume that the fundamental dynamics are also time-symmetric, the only possibility to explain the traces of time asymmetry in experiences lies in the interior condition of (7.21) for empirical predictions. In previous studies, the idea has been discussed that the psychological arrow of time does not require the past hypothesis to be explained, but actually explains the thermodynamic arrow of time [262] (see also [263, 264]). Evaluating this possibility within the framework of the indifference boundary conditions presents an intriguing and significant avenue for future research.

Part III

Matter

Chapter 8

What should be the ontology for the Standard Model?

Although the Standard Model of particle physics is usually formulated in terms of fields, it can be equivalently formulated in terms of particles and strings. In this picture particles and open strings are always coupled. This offers an intuitive and graphical explanation for the otherwise mysterious gauge symmetry. In addition, the particle-string formulation avoids introducing redundant path integral configurations that are present in the field formulation. For its explanatory power and economy, the particle-string ontology may be preferred over the field ontology for the Standard Model.

8.1 Introduction

The Standard Model (of particle physics) is usually formulated as a quantum field theory. A simple-minded understanding of its ontology is in terms of fields. What exists in the Standard Model are field configurations put in superposition under the path integral (Fig. 8.1).

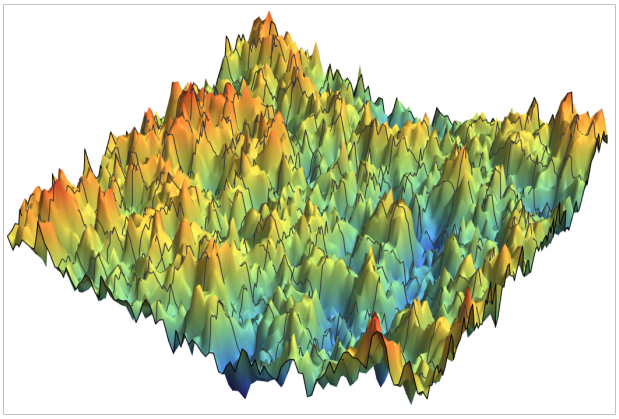

Figure 8.1: The field picture of the Standard Model: What exists are field configurations.

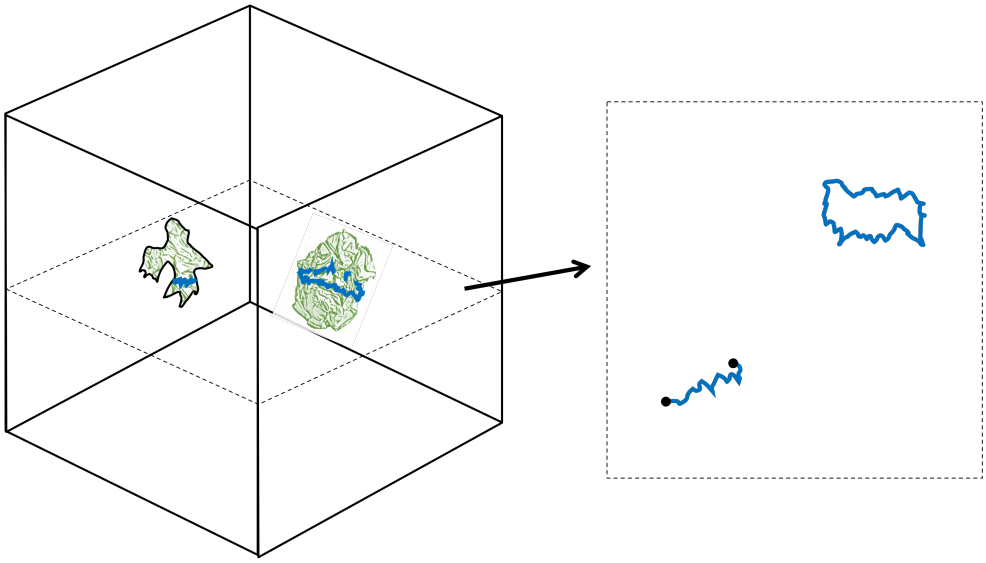

Figure 8.2: The particle-string picture of the Standard Model: What exists are particles and strings, which are generically non-smooth in path integral configurations. In spacetime (left figure), particles trace out $1D$ lines (thick black lines), while strings trace out $2D$ surfaces (green crumpled surfaces). In a hypersurface cross section (right figure), particle lines form points (black points), while string surfaces form lines (blue lines). Crucially, due to gauge symmetry, particles are always attached to strings, and strings are always either closed, or attached to particles.

However, there exists an equivalent reformulation of the Standard Model in terms of particles and strings.[1] In the particle-string formulation[2], what exists are particles and strings bounding each other and put in superposition under a path integral (Fig. 8.2).

The two equivalent formulations pose a question. Is there a preferred ontological picture between the two? Should we think of fields or particles and strings as the basic entities for the Standard Model?

The main point of this work is to point out a conceptual reason to prefer the particle-string picture: It explains gauge symmetries. Suppose the particles are always attached to strings, and suppose the strings are either closed, or have their open ends attached to particles. Then gauge symmetry automatically holds, in the sense that a field reformulation of the particle-string theory automatically obeys gauge symmetries. *In short, the reason for gauge symmetry is that particles and open strings are always coupled.*

In comparison, the field picture leaves gauge symmetries mysterious. There are certainly quantum field theories that are not gauge theories. Why gauge theories then? In the words of Rovelli [278]:

[1] This particle-string formulation not directly related to Superstring theory. Neither extra spacetime dimensions nor supersymmetry is assumed.

[2] Although it is known by many that quantum field theories can generically be re-expressed as theories of particles, strings, and higher-dimensional extended objects at perturbative and non-perturbative levels (see [265, 267, 268, 269, 270, 271, 272, 79, 273] and references therein), this piece of knowledge is not shared by the majority of the physicists. I will follow the works of Gattringer and collaborators [274, 275, 79, 276, 277] to give a review of the particle-string reformulations at the non-perturbative level below. Such reformulations are originally proposed for technical motivations such as more efficient numerical computations. To my knowledge the conceptual implications have not been much investigated. This work is an attempt in this direction.

> Gauge theories are characterized by a local invariance, which is often described as mathematical redundancy. According to this interpretation, physics is coded into the gauge-invariant aspects of the mathematics. ... But things are not so clear. If gauge is only mathematical redundancy, why the common emphasis on the importance of gauge symmetry? Why the idea that this is a major discovery and guiding principle for understanding particle physics?
>
> ... Gauge theories are sometime introduced mentioning the historical idea of promoting a global symmetry to a local one. The purpose of the field would be to realize the local symmetry. This idea, however, leaves the question ... open: why do we need to describe the world with local symmetries if we then interpret these symmetries as mathematical redundancy?

The lack of a convincing explanation for gauge symmetry in the field picture constitutes a reason to prefer the particle-string picture for the ontology of the Standard Model.

The particle-string picture is also more economic than the field picture. For theories with certain global or local symmetries, the field formulation sums over more path integral configurations than the particle-string formulation. These additional configurations cancel among themselves in the end in a field path integral. Avoiding this field redundancy from the outset leads to the particle-string formulation.

An additional motivation to consider the particle-string ontology comes from constructing new theories. For instance in discussions of quantum foundations, Wallace [228] criticizes dynamical collapse models and Bohmian mechanics on the basis that it is much harder than is generally recognised to construct quantum field theory versions of them in order to incorporate the physical contents of the Standard Model. Part of the difficulty is that the dynamical collapse models and Bohmian mechanics studied in the context of non-relativistic quantum physics refer to particles but not fields. One might hope that the particle-string ontology for the Standard Model suggest ways to develop relativistic versions of dynamical collapse models and Bohmian mechanics without the need to migrate to a field ontology.

The above points are elaborated on below. In Section 8.2, I review the particle reformulations of quantum field theories for matter fields. That Z_2 or $U(1)$ global symmetry holds is another way to say that particle lines must keep extending. In Section 8.3, I review the string reformulations of quantum field theories for gauge fields. That $SU(N)$ local symmetry holds is another way to say that strings surfaces must keep extending. In Section 8.4, I review the particle-string reformulation of quantum field theories for matter-gauge coupled systems. That $SU(N)$ local symmetry holds is another way to say that particles and open strings are always coupled. In Section 8.5, I note that quantum field path integrals contain redundant configurations in the presence of the symmetries considered. The redundant configurations are avoided in the particle-string formalism. In Section 8.6, I close with a discussion.

8.2 Matter fields and particles

8.2.1 Real scalar field and unoriented particles

Consider a real scalar field theory in Minkowski spacetime with the Lagrangian density

$$\mathcal{L} = -\frac{1}{2}\partial^\nu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2(x) - V(\phi) \tag{8.1}$$

with a general potential V. To define the path integral non-perturbatively, a D-dimensional hypercubic lattice with spacing a in both time and space directions is introduced. The lattice action reads

$$S = a^D \sum_x [-\frac{1}{2} \sum_{\nu=1}^{D} g^{\nu\nu} (\frac{\phi_{x+\nu} - \phi_x}{a})^2 - \frac{1}{2} m^2 \phi_x^2 - V(\phi_x)] \tag{8.2}$$

$$= \sum_x [\sum_{\nu=1}^{D} g^{\nu\nu} \tilde{\phi}_{x+\nu} \tilde{\phi}_x - \eta \tilde{\phi}_x^2 - \tilde{V}(\tilde{\phi}_x)], \tag{8.3}$$

where $g^{\nu\nu}$ is the Minkowski metric, x refers to lattice vertex, and $x+\nu$ refers to the vertex one unit in the ν direction away from x. In the last line, $\tilde{\phi}_x := a^{\frac{D-2}{2}} \phi_x$, $\eta := a^2 m^2/2 + D - 2$, and $\tilde{V}(\tilde{\phi}_x) = a^D V(\phi_x)$. The tilde symbols are omitted in the following for simplicity.

Particle configurations arise from series expansion

So far there are only field configurations. Particle configurations appear when the path integrand e^{iS} is expressed in its series form. Let S_1 be the first term of (8.3), $\prod_{x,\nu} := \prod_x \prod_{\nu=1}^{D}$, and $\sum_n := \prod_{x,\nu} \sum_{n_{x,\nu}=0}^{\infty}$. Then

$$e^{iS_1} = \prod_{x,\nu} \exp\{i g^{\nu\nu} \phi_{x+\nu} \phi_x\} = \sum_n \prod_{x,\nu} \frac{(i g^{\nu\nu} \phi_{x+\nu} \phi_x)^{n_{x,\nu}}}{n_{x,\nu}!} \tag{8.4}$$

$$= \sum_n (\prod_{x,\nu} \frac{(i g^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!}) (\prod_x \phi_x^{\sum_{\nu=1}^{D} (n_{x,\nu} + n_{x-\nu,\nu})}), \tag{8.5}$$

$$Z = \int D\phi \, e^{iS} = \sum_n (\prod_{x,\nu} \frac{(i g^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!}) (\prod_x \int_{-\infty}^{\infty} d\phi_x \, \phi_x^{n_x} e^{-i\eta\phi_x^2 - iV(\phi_x)}) \tag{8.6}$$

$$= \sum_n (\prod_{x,\nu} \frac{(i g^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!}) (\prod_x f(n_x)), \tag{8.7}$$

where f stands for the integral and $n_x := \sum_{\nu=\pm 1}^{\pm D} n_{x,\nu}$.

These mathematically trivial steps are far from conceptually trivial. The basic entity of the theory has just changed from fields to particles in the following sense. We started with a path integral sum over field values ϕ, but ended with a path integral sum over n in (8.7). Here n assigns a non-negative integer $n_{x,\nu}$ to each lattice edge x, ν connecting x and $x+\nu$. Such an n-configuration admits an interpretation as a particle configuration, with $n_{x,\nu}$ as the number of particles passing the edge x, ν, and n_x as the total number of particle line segments passing x (Fig. 8.3). The path integral of (8.7) is then understood as a sum over particle configurations.

Symmetry and extended particle lines

From (8.6),

$$Z = \sum_n (\prod_{x,\nu} \frac{(i g^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!}) (\prod_x \int_0^{\infty} dr_x \, r_x^{n_x} e^{-i\eta r_x^2} [e^{-iV(r_x)} + (-1)^{n_x} e^{-iV(-r_x)}]), \tag{8.8}$$

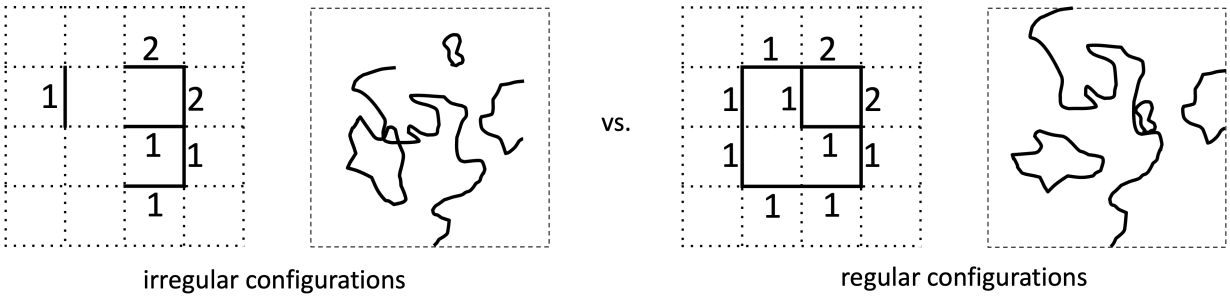

Figure 8.3: Left: lattice and continuum irregular configurations where particle lines do not need to extend (non-zero numbers of particle line segments are labelled on the lattice). Right: lattice and continuum regular configurations where particle lines keep extending.

where ϕ_x is expressed in terms of its magnitude $r_x := |\phi_x|$. When the theory obeys global Z_2 symmetry so that $V(r) = V(-r)$,

$$Z = \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!})(\prod_x \int_0^\infty dr_x \, r_x^{n_x} e^{-i\eta r_x^2}[(1+(-1)^{n_x})e^{-iV(r_x)}]) \tag{8.9}$$

$$= \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu}}}{n_{x,\nu}!})(\prod_x 2\delta_2(n_x) \int_0^\infty dr_x \, r_x^{n_x} e^{-i\eta r_x^2 - iV(r_x)}), \tag{8.10}$$

where $\delta_2(x)$ is the mod 2 Kronecker delta function.

In symmetry considerations it is relevant to draw a distinction between regular and irregular configurations (Fig. 8.3). A regular configuration is such that all the particle lines keep extending until they close on themselves or hit the boundary of the region of spacetime. This requires that in the interior of the region, each particle line segment that enters a vertex is paired with another particle line segment that exits the vertex. In terms of the particle numbers, this is ensured by requiring that at each vertex in the interior of the region, the integers at all the neighboring edges sum to an even value, i.e., by requiring $\delta_2(n_x) = 1$ at all interior vertices x. On the other hand, a configuration that does not obey this requirement is considered irregular, and it contains particle lines that stop extending within the interior of the region.

Before any integration, the path integral (8.6) and equivalently (8.8) include both irregular and regular configurations under $\sum_n$.

For a theory with global Z_2 symmetry, integrating the phase of ϕ (which is ± 1 for a real scalar field) results in $\delta_2(n_x)$. All the irregular configurations cancel out among themselves in the phase sum to leave only the regular configurations where particle lines keep extending. In the particle formulation of the theory, we can declare that the path integral includes only the regular configurations from the outset.

For a theory without global Z_2 symmetry (e.g., with potential $V = \lambda_3\phi^3 + \lambda_4\phi^4$), the irregular configurations are left over. Particle lines can pop up and disappear anywhere in spacetime. Therefore global Z_2 symmetry from the field perspective corresponds to the extendedness of particle lines from the particle perspective.

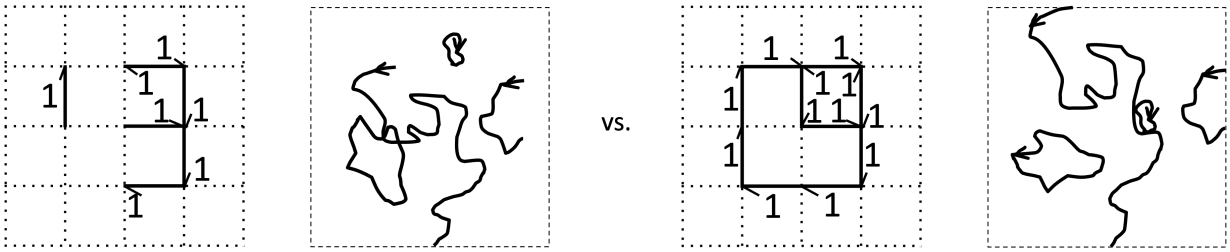

Figure 8.4: Left: lattice and continuum irregular configurations where oriented particle lines do not need to extend (non-zero numbers of particle line segments are labelled on the lattice). Right: lattice and continuum regular configurations where oriented particle lines keep extending.

8.2.2 Complex scalar field and oriented particles

Consider complex scalar field theories in Minkowski spacetime with the Lagrangian density

$$\mathcal{L} = -\partial^\nu \phi \partial_\nu \phi^* - m^2|\phi|^2 - V(\phi). \tag{8.11}$$

The same steps of "non-perturbative definition on lattice-series expansion-imposing symmetry" leads to the following results which are straightforward to derive.

The lattice action reads

$$S = \sum_x [\sum_{\nu=1}^{D} g^{\nu\nu}(\tilde{\phi}_x \tilde{\phi}^*_{x+\nu} + \tilde{\phi}_{x+\nu}\tilde{\phi}^*_x) - \eta\left|\tilde{\phi}_x\right|^2 - \tilde{V}(\tilde{\phi}_x)], \tag{8.12}$$

where $\tilde{\phi}_x := a^{\frac{D-2}{2}}\phi_x$, $\eta := a^2m^2 + 2(D-2)$, and $\tilde{V}(\tilde{\phi}_x) = a^D V(\phi_x)$. The tilde symbols are omitted in the following for simplicity.

Let S_1 be the first term in (8.12), $\prod_{x,\nu} := \prod_x \prod_{\nu=1}^D$, and $\sum_n := \prod_{x,\nu}\sum_{n_{x,\nu}=0}^\infty \sum_{n_{x+\nu,-\nu}=0}^\infty$. Like in the real scalar case, a series expansion leads to [274, 79]

$$e^{iS_1} = \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu}+n_{x+\nu,-\nu}}}{n_{x,\nu}!n_{x+\nu,-\nu}!})(\prod_x \phi_x^{*\sum_\nu(n_{x,\nu}+n_{x,-\nu})}\phi_x^{\sum_\nu(n_{x+\nu,-\nu}+n_{x-\nu,\nu})}), \tag{8.13}$$

$$Z = \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu}+n_{x+\nu,-\nu}}}{n_{x,\nu}!n_{x+\nu,-\nu}!})(\prod_x \int_{-\pi}^{\pi}\frac{d\theta_x}{2\pi} e^{i\theta_x n_x}\int_0^\infty dr_x\, r_x^{\bar{n}_x+1} e^{-i\eta r_x^2 - iV(r_x e^{i\theta_x})}), \tag{8.14}$$

where $\phi \in \mathbb{C}$ is expressed in polar form, and

$$n_x := \sum_{\nu=\pm1}^{\pm D}(n_{x+\nu,-\nu} - n_{x,\nu}), \quad \bar{n}_x := \sum_{\nu=\pm1}^{\pm D}(n_{x+\nu,-\nu} + n_{x,\nu}). \tag{8.15}$$

A lattice edge x,ν is associated with two particle numbers $n_{x,\nu}$ and $n_{x+\nu,-\nu}$ (Fig. 8.4). In the case of real scalar field, we interpreted the one integer on an edge as the unoriented particle number. Here we interpret the two integers as the numbers of oriented particles travelling the edge in different directions. In particular, $n_{x,\nu}$ represents the number of particles travelling from

x in the positive ν direction, while $n_{x+\nu,-\nu}$ represents the number of particles travelling from $x+\nu$ in the negative ν direction. Then n_x represents the difference between the numbers of particle line segments entering and exiting x, and $\bar{n}_x$ represents the total number of particle line segments crossing x.

Again, in symmetry considerations it is relevant to draw a distinction between regular and irregular configurations (Fig. 8.4). A regular configuration is such that all the oriented particle lines keep extending until they close on themselves or hit the boundary of the region of spacetime. This requires that in the interior of the region, each particle line segment that enters a vertex is paired with another particle line segment that exits the vertex. Since n_x represents the difference between the numbers of particle line segments entering and exiting x, this is ensured by requiring $\delta(n_x) = 1$ at all interior vertices x. On the other hand, a configuration that does not obey this requirement is considered irregular, and it contains particle lines that stop extending within the interior of the region.

Before any integration, the path integral (8.14) includes both irregular and regular configurations under $\sum_n$.

For a theory with global $U(1)$ symmetry so that $V(re^{i\theta}) = V(r)$,

$$Z = \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu}+n_{x+\nu,-\nu}}}{n_{x,\nu}!n_{x+\nu,-\nu}!})(\prod_x \delta(n_x) \int_0^\infty dr_x \, r_x^{\bar{n}_x+1} e^{-i\eta r_x^2 - iV(r_x)}). \tag{8.16}$$

Integrating the phase of ϕ results in $\delta(n_x)$, which implies that the number of incoming and outgoing line segments are equal at all vertices. All the irregular configurations cancel out among themselves in the phase integral to leave only the regular configurations where particle lines keep extending (Fig. 8.4). In the particle formulation of the theory, we can declare that the path integral includes only the regular configurations from the outset.

For a theory without global $U(1)$ symmetry (e.g., with potential $V = \lambda_3\phi^3 + \lambda_4|\phi|^4$), the irregular configurations are left over. Particle lines can pop up and disappear anywhere in spacetime. Therefore global $U(1)$ symmetry from the field perspective corresponds to the extendedness of oriented particle lines from the particle perspective.

8.2.3 Fermion field and oriented particles

For a fermionic theory in Minkowski spacetime with the Lagrangian density

$$\mathcal{L} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{8.17}$$

the lattice action reads (after redefinitions to absorb constants)[3]

$$S = \sum_x [\sum_{\mu=1}^{D} (\overline{\psi}_x i\gamma^\mu \psi_{x+\mu} - \overline{\psi}_{x+\mu} i\gamma^\mu \psi_x) - m\overline{\psi}_x\psi_x]. \tag{8.18}$$

With $\prod_x$ and $\prod_{x,\mu}$ as defined previously, a series expansion yields

$$e^{iS} = [\prod_x \sum_{s_x} (-im\overline{\psi}_x\psi_x)^{s_x}] \prod_{x,\mu} [\sum_{n_{x,\mu}} (-\overline{\psi}_x\gamma^\mu\psi_{x+\mu})^{n_{x,\mu}} \sum_{n_{x+\mu,-\mu}} (\overline{\psi}_{x+\mu}\gamma^\mu\psi_x)^{n_{x+\mu,-\mu}}], \tag{8.19}$$

[3] In practical studies of lattice field theory one can adopt alternative actions such as with staggered fermions [276, 277] to ameliorate the fermion doubling problem. From a fundamental perspective perhaps it is more satisfactory to stick to the less *ad hoc* naive fermion action and subject lattice lengths to path integration in quantum gravity [18].

where $s, n = 0, 1$ because Grassmann variables are nilpotent. For the partition function, Grassmann integration yields

$$Z = \int D[\overline{\psi}, \psi][\prod_x \sum_{s_x} (-im\overline{\psi}_x\psi_x)^{s_x}] \prod_{x,\mu} [\sum_{n_{x,\mu}} (-\overline{\psi}_x\gamma^\mu\psi_{x+\mu})^{n_{x,\mu}} \sum_{n_{x+\mu,-\mu}} (\overline{\psi}_{x+\mu}\gamma^\mu\psi_x)^{n_{x+\mu,-\mu}}] \quad (8.20)$$

$$= \sum_{s,n} (-im)^{s_x} (-\gamma^\mu)^{n_{x,\mu}} (\gamma^\mu)^{n_{x+\mu,-\mu}} \prod_x \delta(u_x)\delta(v_x) \quad (8.21)$$

$$= C \sum_n (\frac{\gamma^\mu}{im})^{n_{x,\mu}} (-\frac{\gamma^\mu}{im})^{n_{x+\mu,-\mu}} \prod_x C_x(n). \quad (8.22)$$

In the second line, $u_x := s_x + \sum_{\mu=\pm 1}^D n_{x,\mu} - 1$, $v_x := s_x + \sum_{\mu=\pm 1}^D n_{x+\mu,-\mu} - 1$, and the delta functions are induced by Grassmann integration. Pictorially, in each configuration a lattice site is either filled by a mass "monomer" counted by s, or is crossed by exactly one outgoing and exactly one incoming fermion line segment counted by n.

In the third line a factor $-im$ is extracted from each site to form $C = (-im)^N$ where N is the number of sites. For sites occupied by a monomer, this comes from the monomer factor. For sites on a fermion line, this induces a division by $-im$, which can be attributed to the γ^μ factors, since the number of sites on a fermion line equals the number of line segments on it (away from the boundary of the region). After the s-sum, the constraint $C_x(n)$ enforces that a site x is crossed by either 0 or 1 fermion line. Explicitly

$$C_x(n) = \delta(n_x)[\delta(\bar{n}_x) + \delta(\bar{n}_x - 2)], \quad (8.23)$$

where n_x and $\bar{n}_x$ are as defined in (8.15). They represent the difference in number for incoming and outgoing line segments, and the total number of line segments crossing x.

This picture of the extending oriented fermion particles is the same as that of the complex scalar field with $U(1)$ symmetry, except that no identical fermion line segments can overlap which enforces Pauli's exclusion principle. For fermions we do not consider theories with non-extending particle lines (irregular configurations) because any Lagrangian density where ψ and $\overline{\psi}$ show up in pairs automatically enforces global $U(1)$ symmetry.

8.3 Gauge fields and strings

Quantum field theory presentations of gauge theories sometimes leave the impression that gauge matter is not much different from scalar and fermion matter. Gauge field, like scalar and fermion fields, is just another field, with perhaps more components.

In contrast, the particle-string formulation makes it clear that gauge matter is a totally different species. While the scalar and spin-$1/2$ fermion matter are particles tracing out $1D$ lines in spacetime, gauge matter are strings tracing out $2D$ surfaces in spacetime.

8.3.1 Abelian gauge field and oriented surfaces

The standard way to define quantum gauge theories is through Wilson's lattice gauge theory formalism [267]. For a $U(1)$ gauge field with the Lagrangian density $\mathcal{L} = \frac{\beta}{2} F_{\mu\nu}^2$, the lattice action

in terms of group variables $U_{x,\mu} \in U(1)$ on edges is

$$S = \frac{\beta}{2} \sum_{x,\mu<\nu} g^{\mu\mu} g^{\nu\nu} (U_{x,\mu} U_{x+\mu,\nu} U^*_{x+\nu,\mu} U^*_{x,\nu} + U^*_{x,\mu} U^*_{x+\mu,\nu} U_{x+\nu,\mu} U_{x,\nu}), \tag{8.24}$$

where $\prod_{x,\mu<\nu} := \prod_x \prod_{\nu=2}^{D} \prod_{\mu=1}^{\nu-1}$., which is a sum over all plaquettes (elementary surfaces) of the lattice.

String configurations arise from series expansion

Let I_n be the modified Bessel function defined by $e^{\frac{z}{2}(t+t^{-1})} = \sum_{n\in\mathbb{Z}} I_n(z) t^n$ for $z, t \in \mathbb{C}, t \neq 0$. Then [79]

$$Z = \int D[U] e^{\frac{i\beta}{2} \sum_{x,\mu<\nu} g^{\mu\mu} g^{\nu\nu} (U_{x,\mu} U_{x+\mu,\nu} U^*_{x+\nu,\mu} U^*_{x,\nu} + U^*_{x,\mu} U^*_{x+\mu,\nu} U_{x+\nu,\mu} U_{x,\nu})} \tag{8.25}$$

$$= \int D[U] \prod_{x,\mu<\nu} \sum_{p_{x,\mu\nu}\in\mathbb{Z}} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu} g^{\nu\nu}] (U_{x,\mu} U_{x+\mu,\nu} U^*_{x+\nu,\mu} U^*_{x,\nu})^{p_{x,\mu\nu}} \tag{8.26}$$

$$= \sum_p \Big(\prod_{x,\mu<\nu} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu} g^{\nu\nu}] \Big) \prod_{x,\mu} \int dU_{x,\mu} U^{p_{x,\mu}}_{x,\mu} \tag{8.27}$$

$$= \sum_p \Big(\prod_{x,\mu<\nu} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu} g^{\nu\nu}] \Big) \prod_{x,\mu} \delta(p_{x,\mu}), \tag{8.28}$$

where $\sum_p := \prod_{x,\mu<\nu} \sum_{p_{x,\mu\nu}\in\mathbb{Z}}$, $\prod_{x,\mu} := \prod_x \prod_{\mu=1}^{D}$, and

$$p_{x,\mu} := \sum_{\rho:\rho<\mu} (p_{x,\rho\mu} - p_{x-\rho,\rho\mu}) - \sum_{\rho:\rho>\mu} (p_{x,\mu\rho} - p_{x-\rho,\mu\rho}). \tag{8.29}$$

We started with a path integral sum over field values U, but ended with a path integral sum over p in (8.28). Here p assigns an integer $p_{x,\mu\nu} \in \mathbb{Z}$ to each lattice plaquette $x, \mu\nu$ starting at vertex x and extending in directions μ and ν. Such a p-configuration admits an interpretation as a string configuration, with $p_{x,\mu\nu}$ as the number of elementary surface segments at the lattice plaquette $x, \mu\nu$, positively or negatively orientated depending on the sign of $p_{x,\mu\nu} \in \mathbb{Z}$. The path integral of (8.28) is then understood as a sum over string configurations.

Symmetry and extended string surfaces

There is again a distinction between regular and irregular configurations (Fig. 8.5). A regular configuration is such that all the oriented surfaces keep extending until they close on themselves or hit the boundary of the region of spacetime. According to (8.29), the net number (positive minus negative oriented) of surface segments crossing the edge x, μ is $p_{x,\mu}$. Requiring that the oriented surfaces always extend across the edge amounts to demanding $\delta(p_{x,\mu}) = 0$ at all interior edges, because this means all elementary surfaces touching this edge can be glued each other in a consistent orientation. On the other hand, a configuration that does not obey this requirement is considered irregular, and it contains string surfaces that stop extending within the interior of the region.

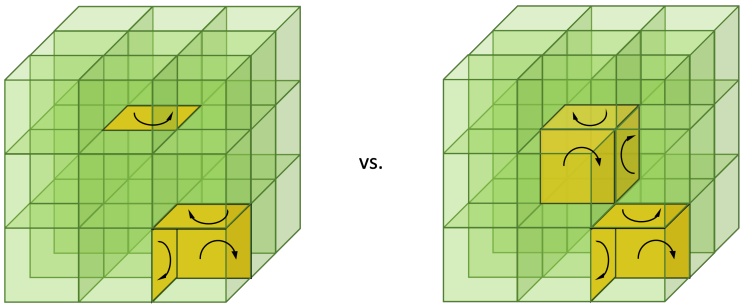

Figure 8.5: Left: an irregular surface configuration where positively and negatively oriented surface numbers do not match on some edges. Right: a regular surface configuration where positively and negatively oriented surface numbers match on all edges (in the interior of the region under consideration).

Before any integration, the path integral (8.27) include both irregular and regular configurations under $\sum_p$.

Integrating over U in (8.28) gives rise to the delta function $\delta(p_{x,\mu})$, which ensure that the irregular configurations cancel out, and that only regular string configurations appear in the sum $\sum_p$. More generally, when a theory obeys gauge symmetry, the action contains only terms such as $U_{x,\mu}U_{x+\mu,\nu}U^*_{x+\nu,\mu}U^*_{x,\nu}$ where the edge variables U form closed loops. A series expansion leads to a polynomial in the loops, which upon integration by $\int D[U]$ generates the delta function. In the surface picture, this implies that the oriented surfaces keep extending.

In contrast, suppose we start with a more general action $S = \int d^4x[-\frac{\beta}{2}F^2_{\mu\nu} + V(A_\mu)]$, such as the Proca action with $V(A_\mu) = m^2 A_\mu A^\mu$. Then the non-perturbative theory has to resort from $U_{x,\mu} = e^{iaA_{x,\mu}}$ to $A_{x,\mu}$ as the basic variable in order to accommodate the additional term $V(A_\mu)$ in the action. The same procedure as in (8.25) to (8.27) yields

$$Z = \int D[A] e^{\frac{i\beta}{2}\sum_{x,\mu<\nu} g^{\mu\mu}g^{\nu\nu}(U_{x,\mu}U_{x+\mu,\nu}U^*_{x+\nu,\mu}U^*_{x,\nu} + U^*_{x,\mu}U^*_{x+\mu,\nu}U_{x+\nu,\mu}U_{x,\nu})} e^{i\sum_x V(A_{x,\mu})} \tag{8.30}$$

$$= \sum_p \Big(\prod_{x,\mu<\nu} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu}g^{\nu\nu}]\Big) \int \Big(\prod_{x,\mu} DA_{x,\mu}\Big)\, e^{ia\sum_{x,\mu} A_{x,\mu}p_{x,\mu} + i\sum_x V(A_{x,\mu})}. \tag{8.31}$$

For a general $V(A_{x,\mu})$ the constraint $\delta(p_{x,\mu})$ can no longer be derived. The surface picture where $\sum_p$ represents a sum over string configurations still holds, but the surfaces no longer need to keep extending.

8.3.2 Non-Abelian gauge field and colored oriented surfaces

As an example of a non-Abelian gauge theory, consider $SU(3)$ gauge theory with the lattice action

$$S = \frac{\beta}{6}\sum_{x,\mu<\nu}\sum_{a,b,c,d=1}^{3} g^{\mu\mu}g^{\nu\nu}(U^{ab}_{x,\mu}U^{bc}_{x+\mu,\nu}U^{dc\,*}_{x+\nu,\mu}U^{ad\,*}_{x,\nu} + U^{ab\,*}_{x,\mu}U^{bc\,*}_{x+\mu,\nu}U^{dc}_{x+\nu,\mu}U^{ad}_{x,\nu}), \tag{8.32}$$

where U^{ab} is the matrix representation of the $SU(3)$ group element in some basis. Similar to the $U(1)$ case, expanding the path integrand as a series in U^{ab} and performing the group integration

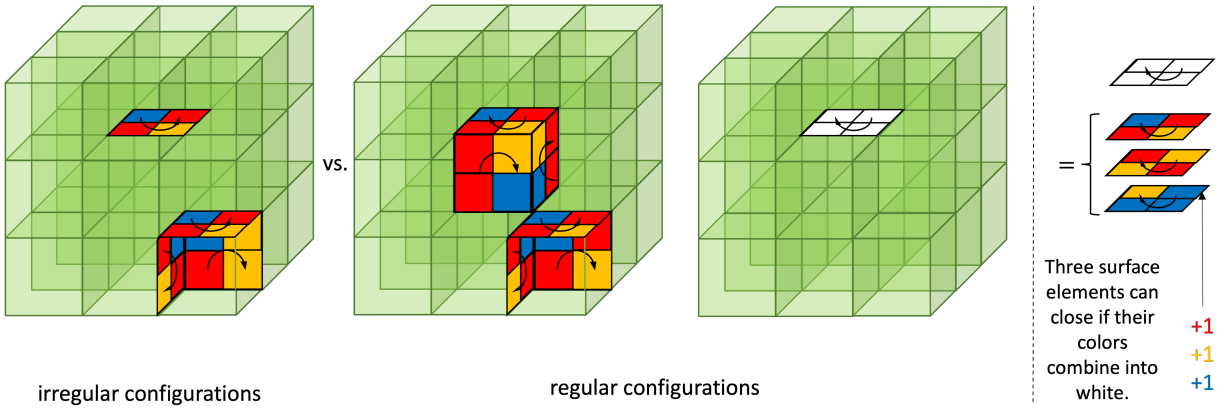

Figure 8.6: Left: an irregular surface configuration where either color or positively and negatively oriented surface numbers do not match on some edges. Middle: a regular surface configuration where both color and positively and negatively oriented surface numbers match on all edges (in the interior of the region under consideration). Right: a regular surface configuration closed by color combination.

in terms of the various matrix components U^{ab} yields delta function constraints. The explicit formulas which can be found in [276, 277] will not be shown here, because they are a bit lengthy and are not used below.

Again, a picture of extended string surfaces (Fig. 8.6) arises that captures the essence of the result. On each edge the 3×3 matrix elements $U^{ab}_{x,\mu}$ introduces 3×3 color combinations, e.g., red into yellow, red into blue etc. A lattice plaquette then has four color slots to support 3^4 color combinations for a surface element.

The constraints arising from group integration indicate that only regular configurations where surfaces keep extending stay (Fig. 8.6). Here a surface element extends by matching in color and cancelling in orientation (positive orientation cancels with negative orientation) on the common edge(s) with an adjacent surface element. This part is a straightforward generalization of the $U(1)$ case with one color.

Interestingly, in $SU(3)$ surfaces can also close through color combination. Suppose three surface elements overlap in the same direction on an edge. At one of the two slots, if all three colors are present they combine into white. If the color is white at both slots of the edge, the three surfaces are considered to have no boundary at this edge. Through color combination, new closed surfaces can form, such as the totally white surface on the right of Fig. 8.6.

Again, the constraints are generated out of group integration. If we start with a theory without non-Abelian gauge symmetry, the constraints can no longer be derived. Pictorially, colored oriented surfaces no longer need to keep extending.

8.4 Particle-string coupling

Consider a theory with both particles and strings. Suppose the theory obeys gauge symmetry (in the field reformulation). Then as demonstrated below, the particles are always attached to the

strings, and the strings are either closed, or have their open ends attached to the particles. This need not hold if the theory does not obey gauge symmetry.

If one adopts the particle-string ontology, then an intuitive explanation for gauge symmetry is available. Suppose the particles and open strings are always coupled. Then the field reformulation for the theory automatically obeys gauge symmetry. In this view, gauge symmetry is no longer a guiding principle, but only a derived property.

8.4.1 Particles and uncolored strings

Consider a $U(1)$-locally symmetric scalar-gauge coupled theory with the Lagrangian density

$$\mathcal{L} = -|D_\mu \phi|^2 - m^2|\phi|^2 + V(\phi_x) + \frac{\beta}{2} F_{\mu\nu}^2. \tag{8.33}$$

where $D_\mu = \partial_\mu + igA_\mu$ and V is a $U(1)$-locally symmetric potential. Non-perturbatively, the scalar part action changes to

$$S_P = \sum_x [\sum_{\nu=1}^{D} g^{\nu\nu}(\phi_x U_{x,\nu}^* \phi_{x+\nu}^* + \phi_{x+\nu} U_{x,\nu} \phi_x^*) - \eta |\phi_x|^2 - V(\phi_x)], \tag{8.34}$$

where suitable redefinitions are performed as in Section 8.2.2. In comparison to (8.12), the only difference is that every near neighbor coupling $\phi_x \phi_{x+\nu}^*$ is now dressed with $U_{x,\nu}^*$. After the ϕ integration, the same steps as in Section 8.2.2 leads to

$$Z_P = \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu} U_{x,\nu})^{n_{x,\nu}} (ig^{\nu\nu} U_{x,\nu}^*)^{n_{x+\nu,-\nu}}}{n_{x,\nu}! n_{x+\nu,-\nu}!}) (\prod_x \delta(n_x) \int_0^\infty dr_x \, r_x^{\bar{n}_x + 1} e^{-i\eta r_x^2 - iV(r_x)}) \tag{8.35}$$

for the particle part. As in Section 8.2.2, we interpret the n-configurations as oriented particle configurations. Due to the delta function, the extended particle lines picture persists for the particle part of this $U(1)$-locally symmetric theory.

The gauge part of the theory is the same as in Section 8.3.1, except for the U factors coming from the dresses on the particle lines. Therefore after the U integration, (8.28) is replaced by

$$\sum_p (\prod_{x,\mu<\nu} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu} g^{\nu\nu}]) \prod_{x,\mu} \delta(p_{x,\mu} + n_{x,\mu}). \tag{8.36}$$

All together,

$$\begin{aligned} Z = &\sum_p \sum_n (\prod_{x,\nu} \frac{(ig^{\nu\nu})^{n_{x,\nu} + n_{x+\nu,-\nu}}}{n_{x,\nu}! n_{x+\nu,-\nu}!}) (\prod_x \delta(n_x) \int_0^\infty dr_x \, r_x^{\bar{n}_x + 1} e^{-i\eta r_x^2 - iV(r_x)}) + \\ &(\prod_{x,\mu<\nu} I_{p_{x,\mu\nu}}[i\beta g^{\mu\mu} g^{\nu\nu}]) \prod_{x,\mu} \delta(p_{x,\mu} + n_{x,\mu}). \end{aligned} \tag{8.37}$$

As in Section 8.3.1, we interpret the p-configurations as oriented string configurations, with $p_{x,\mu\nu} \in \mathbb{Z}$ as the number of elementary surface segments at the lattice plaquette $x, \mu\nu$, positively or negatively orientated depending on the sign of $p_{x,\mu\nu}$.

The delta function $\delta(p_{x,\mu} + n_{x,\mu})$ of (8.37) indicates that the number of positively or negatively oriented elementary particle lines matches with the number of negatively or positively oriented

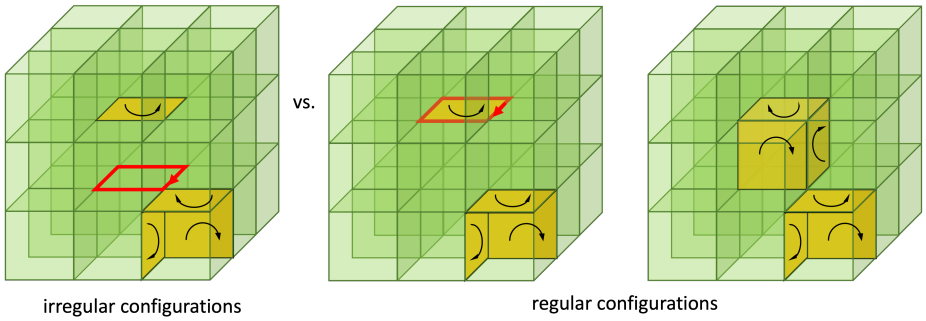

Figure 8.7: Left: an irregular line-surface configuration where positively and negatively oriented line and surface numbers do not match on some edges. Right: regular surface configurations where positively and negatively oriented line and surface numbers match on all edges (in the interior of the region under consideration).

elementary surfaces at the edges x, μ. Pictorially, this admits the interpretation that particle lines are always attached to string surfaces, and the open boundaries of string surfaces are always bounded by particle lines. In addition, the orientations are always matched (Fig. 8.7). Referring to configurations obeying these conditions as regular configurations and those that do not as irregular configurations, we can declare that in the particle-string formulation of the theory, the path integral defined by (8.37) includes only the regular configurations from the outset.

If we started with a theory without gauge symmetry, this picture no longer holds. For instance, suppose

$$V(\phi_x) = \sum_{\nu=1}^{D} g^{\nu\nu}(\phi_x \phi^*_{x+\nu} + \phi_{x+\nu}\phi^*_x) \tag{8.38}$$

is as in a bare scalar action without gauge coupling. Then in place of (8.34) and (8.35) we would get

$$S_P = \sum_x [\sum_{\nu=1}^{D} g^{\nu\nu}[\phi_x(U^*_{x,\nu}+1)\phi^*_{x+\nu} + \phi_{x+\nu}(U_{x,\nu}+1)\phi^*_x] - \eta|\phi_x|^2], \tag{8.39}$$

$$Z_P = \sum_n (\prod_{x,\nu} \frac{[ig^{\nu\nu}(U_{x,\nu}+1)]^{n_{x,\nu}}[ig^{\nu\nu}(U^*_{x,\nu}+1)]^{n_{x+\nu,-\nu}}}{n_{x,\nu}! n_{x+\nu,-\nu}!})(\prod_x \delta(n_x) \int_0^\infty dr_x \, r_x^{\bar{n}_x+1} e^{-i\eta r_x^2}). \tag{8.40}$$

Since U now shows up as $(U^*_{x,\nu}+1)$ and $(U^*_{x,\nu}+1)$,the integrals over U on the edges would no longer yield delta functions $\delta(p_{x,\mu}+n_{x,\mu})$ on the edges. This non-gauge invariant theory therefore include irregular configurations where particle lines and open string surfaces are not attached to each other. This example shows why the gauge-invariant coupling

$$g^{\nu\nu}(\phi_x U^*_{x,\nu}\phi^*_{x+\nu} + \phi_{x+\nu}U_{x,\nu}\phi^*_x) \tag{8.41}$$

for ϕ on adjacent vertices is crucial for generating the delta functions $\delta(p_{x,\mu}+n_{x,\mu})$ on the edges to exclude irregular configurations.

For fermion particles coupled to strings, apart from the fact that identical fermionic lines segments cannot overlap (Pauli's exclusion principle), the picture is the same.

8.4.2 Particles and colored strings

For a non-Abelian locally symmetric scalar-gauge coupled theory, the gauge part of the action is as (8.32), and the scalar part of the action is

$$S_P = \sum_x [\sum_{\nu=1}^{D} \sum_{a,b} g^{\nu\nu} (\phi_x^a U_{x,\nu}^{ab\,*} \phi_{x+\nu}^{b\,*} + \phi_{x+\nu}^a U_{x,\nu}^{ab} \phi_x^{b\,*}) - \eta |\phi_x^a|^2 - V(\phi_x^a)], \tag{8.42}$$

where the sum $\sum_a$ is over colors.

The same steps as in the last example leads to a picture of colored particles coupled to colored strings. Since a particle line segment is now dressed with U^{ab}, it also carries two colors. Group integration implies that the colors of the particle lines and of the string surfaces must cancel. In addition to the closed string surfaces shown in Section 8.3.2, there are now open string surfaces bounded by particle lines, whose colors and orientations match.

Again, the picture of the fermion particles [276, 277] is quite the same apart from the fact that identical fermionic lines cannot overlap (Pauli's exclusion principle).

8.5 Field redundancy and partial local symmetry

8.5.1 Field redundancy

In the previous examples, quantum field theories with symmetries exhibit a redundancy. The irregular configurations cancel out among all themselves in the path integral sum. Only the regular configurations need to be included from the very beginning.

In terms of particles and strings, the path integral can be defined to include only regular configurations. The redundancy is avoided.

In contrast, in terms of fields the irregular configurations seems unavoidable because the form of quantum field theories is tightly constrained [279]. The particle-string formalism is more economic than the field formalism in this regard.

Because this form of redundancy is attached to the field formalism and can be avoided in the particle-string formalism, I call it **field redundancy**. Field redundancy is distinct from the gauge redundancy that relates gauge equivalent configurations. In the previous examples, field redundancy is seen for a broader class of symmetries including discrete, continuous, local, global, Abelian, and non-Abelian symmetries.

Field redundancy is also a quantum property foreign to classical theory, because the cancellation is among configurations in superposition under a path integral.

8.5.2 Partial local symmetries

As an aside question, is there a more precise way to capture the relation of field redundancy to symmetry?

It is tempting to understand the cancellation in terms of Noether's theorem that relates charge conservation to symmetry, because particles and strings that keep extending seem to suggest some

form of conservation law. However, Noether's theorem cannot be the answer. Firstly, Noether's theorem does not cover discrete symmetries such as the Z_2 symmetry of the real scalar field. Secondly, non-Abelian gauge theories do not have gauge-invariant Noether currents [280], whereas the regular particle-string configurations here are gauge invariant configurations.

The true answer seems to be given by what I call **partial local symmetry**. Consider path integrals that can be re-expressed as follows.

$$\begin{aligned} Z &= \int D[Y]\ A(Y) \\ &= \int D[R](\prod_x D[G_x])\ A(R,\{G_x\}) \\ &= \int D[R]\ A_0(R) \prod_x \underbrace{(\int D[G_x]\ P_x(R,G_x))}_{C_x}\,. \end{aligned}$$

In the second line, the original variable Y is decomposed into group variables $\{G_x\}$ and residue variable(s) R, and the amplitude $A(Y)$ is re-expressed in these new variables as $A(R,\{G_x\})$. Here the group variables G_x are located to places x, which can be lattice points, edges, plaquettes etc.

In the third line, $A(R,\{G_x\}) = A_0(R)(\prod_x A_x(R,G_x))$ decomposes into two parts. The $A_0(R)$ part is independent of the group variables so can be taken out of the group integrals. This part is invariant under the local group actions by G_x at any location x. Therefore the theory exhibits a form of local symmetry. Since $A_0(R)$ is only part of the whole amplitude $A(Y)$, the symmetry is dubbed "partial local symmetry".

The other part $P_x(R,G_x)$ are polynomials in the group variables G_x. Group integration $\int D[G_x]\ A_x(R,G_x)$ generates local constraints C_x for the residue variables R.

For example, for the Z_2-symmetric real scalar field theory (8.9), $C_x = \sum_{z=\pm 1} z^{n_x} = \delta_2(n_x)$ with $G_x \in Z_2$, $R = n_x$, and $P_x = G_x^{n_x}$. As another example, for the $U(1)$-symmetric complex scalar field theory (8.16), $C_x = \int_{-\pi}^{\pi} \frac{d\theta_x}{2\pi} e^{i\theta_x n_x} = \delta(n_x)$ with $G_x \in U(1)$, $R = n_x$, and $P_x = G_x^{n_x}$. The other examples in the previous sections also follow the same pattern.

Although these theories do not necessarily exhibit local gauge symmetry, they do exhibit partial local symmetry. The constraints that arise result in cancellations among field configurations, and hence imply field redundancy.

8.6 Discussion

Because the Standard Model is commonly formulated as a quantum field theory, it is tempting to consider fields as its basic entity, i.e., its ontology. In light of the particle-string reformulation reviewed here[4], the question about the ontology of the Standard Model deserves a deeper thought.

In the particle-string formulation, particles and open strings are always coupled. This property gives an intuitive explanation for the otherwise mysterious gauge symmetry of the field formulation. In addition, the field formulation includes redundant configurations that eventually

[4]To be precise, the particle-string formulation of the full Standard Model is not explicitly given, but it can be straightforwardly obtained by generalizing the particle-string formulation of the matter-gauge coupled theories.

cancel out among themselves. These redundant configurations are avoided from the outset in the particle-string formulation. For its explanatory power and its economy, the particle-string ontology may be preferred over the field ontology for the Standard Model.

One could question if the particle-string formulation really explains gauge symmetry better. After all, *a priori* particles and open strings need not always be coupled, and one still needs to assume that they are in order to explain gauge symmetry. If it costs an extra assumption anyways, why could we not simply assume that the gauge fields obey a local symmetry to explain gauge symmetry?

I believe the particle-string explanation is still preferable. Imagine we are to explain the fundamental properties of matter to school students or laypeople eager for the scientific knowledge. The gauge symmetry principle is such an profound property that we do not want to miss. One way to explain it is: "Matter are made of fields. The fields can be transformed according to local group actions. In our universe, the dynamical laws for matter are unchanged under such transformations." Another way is: "Matter are made of particles and strings. In our universe, particles and open strings are always coupled." Which one do you prefer? An explanation is supposed to build intuition. The former "explanation" hardly builds any intuition at all, and should perhaps better be characterized as a "description" instead of an "explanation" for gauge symmetry. The latter explanation reduces the complicated mathematical concept of gauge symmetry to an easily visualizable picture of coupled particle lines and string surfaces. It does help build intuition, and is in this sense preferable.

These in no way implies that we should abandon the field formulation. A calculation that is hard in one formulation can be easier in another. For practical uses it is better to have more formulations in our toolbox than fewer, even though for conceptual understandings one formulation may be preferred.

In addition to these, reformulating a known theory in another picture can suggest different ideas towards discovering the unknowns. It is worth exploring ideas of beyond the Standard Model coming from the particle-string picture. For instance, a $4D$ space time has room to support higher dimensional objects in addition to $1D$ particle lines and $2D$ string surfaces. Could dark matter be such higher dimensional objects?

References

[1] Ding Jia. What Should Be the Ontology for the Standard Model? *Foundations of Physics*, 52(4):1–20, 8 2022.

[2] Ding Jia. Complex, Lorentzian, and Euclidean simplicial quantum gravity: numerical methods and physical prospects. *Classical and Quantum Gravity*, 39(6):065002, 2 2022.

[3] Ding Jia. Light ray fluctuations in simplicial quantum gravity. *Classical and Quantum Gravity*, 39(13):135005, 6 2022.

[4] Ding Jia. Truly Lorentzian quantum cosmology. *arXiv: 2211.00517*.

[5] Ding Jia. Time-space duality in 2D quantum gravity. *Classical and Quantum Gravity*, 39(3):035016, 1 2022.

[6] Ding Jia. Light ray fluctuation and lattice refinement of simplicial quantum gravity. *Classical and Quantum Gravity*, 40(16):165003, 7 2023.

[7] Ding Jia. Quantum gravity and time order. *(unpublished notes)*.

[8] Ding Jia. Is singularity resolution trivial? *arXiv: 2204.12304*.

[9] Ding Jia. Experience in quantum physics: toward a theory of everything. *arXiv:2306.11549*.

[10] Ding Jia. Modes of experience in a superposed world. *arXiv:2208.10920*.

[11] Ding Jia. Decoherence does not imply branching. *(to appear)*.

[12] Ding Jia. Path integral and particle ontology. *arXiv: 2209.05345*.

[13] Richard P. Feynman and Albert R. Hibbs. *Quantum Mechanics and Path Integrals*. McGraw-Hill, New York, 1965.

[14] T. Regge. General relativity without coordinates. *Il Nuovo Cimento Series 10*, 19(3):558–571, 2 1961.

[15] M. Roček and Ruth M. Williams. Quantum Regge calculus. *Physics Letters B*, 104(1):31–37, 8 1981.

[16] R. M. Williams and P. A. Tuckey. Regge calculus: A brief review and bibliography. *Classical and Quantum Gravity*, 9(5):1409–1422, 5 1992.

[17] Renate Loll. Discrete Approaches to Quantum Gravity in Four Dimensions. *Living Reviews in Relativity*, 1(1):13, 12 1998.

[18] H. W. Hamber. *Quantum Gravitation : the Feynman Path Integral Approach*. Springer, Berlin, 2009.

[19] John W. Barrett, Daniele Oriti, and Ruth M. Williams. Tullio Regge's legacy: Regge calculus and discrete gravity. In Leonardo Castellani, Anna Ceresole, Riccardo D'Auria, and Pietro Fré, editors, *Tullio Regge: An Eclectic Genius*. World Scientific, Singapore, 2019.

[20] G. W. Gibbons. The Einstein action of Riemannian metrics and its relation to quantum gravity and thermodynamics. *Physics Letters A*, 61(1):3–5, 4 1977.

[21] Jan Ambjørn, Jakob L Nielsen, Juri Rolf, and George Savvidy. Spikes in quantum Regge calculus. *Classical and Quantum Gravity*, 14(12):3225, 1997.

[22] Herbert Hamber. Vacuum Condensate Picture of Quantum Gravity. *Symmetry*, 11(1):87, 1 2019.

[23] Kyle Tate and Matt Visser. Fixed-topology Lorentzian triangulations: Quantum Regge Calculus in the Lorentzian domain. *Journal of High Energy Physics*, 2011(11):72, 11 2011.

[24] Kyle Tate and Matt Visser. Realizability of the Lorentzian (n, 1)-simplex. *Journal of High Energy Physics*, 2012(1):28, 1 2012.

[25] Aleksandar Mikovic. Piecewise Flat Metrics and Quantum Gravity. *arXiv: 2001.11439*.

[26] Seth K. Asante, Bianca Dittrich, and José Padua-Argüelles. Effective spin foam models for Lorentzian quantum gravity. *Classical and Quantum Gravity*, 38(19):195002, 9 2021.

[27] Bianca Dittrich, Steffen Gielen, and Susanne Schander. Lorentzian quantum cosmology goes simplicial. *Classical and Quantum Gravity*, 39(3):035012, 1 2022.

[28] J. Ambjorn, A. Goerlich, J. Jurkiewicz, and R. Loll. Nonperturbative Quantum Gravity. *Physics Reports*, 519(4-5):127–210, 3 2012.

[29] R. Loll. Quantum gravity from causal dynamical triangulations: A review. *Classical and Quantum Gravity*, 37(1):013002, 12 2020.

[30] AuroraScience Collaboration, Marco Cristoforetti, Francesco Di Renzo, and Luigi Scorzato. High density QCD on a Lefschetz thimble? *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 86(7), 5 2012.

[31] Andrei Alexandru, Gökçe Başar, Paulo F. Bedaque, and Neill C. Warrington. Complex paths around the sign problem. *Reviews of Modern Physics*, 94(1):015006, 3 2022.

[32] Muxin Han, Zichang Huang, Hongguang Liu, Dongxue Qu, and Yidun Wan. Spinfoam on a Lefschetz thimble: Markov chain Monte Carlo computation of a Lorentzian spinfoam propagator. *Physical Review D*, 103(8):084026, 4 2021.

[33] Andrei Alexandru, Gökçe Basar, Paulo F. Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Sign problem and Monte Carlo calculations beyond Lefschetz thimbles. *Journal of High Energy Physics 2016 2016:5*, 2016(5):1–17, 5 2016.

[34] Andrei Alexandru, Gökçe Başar, Paulo F. Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Monte Carlo calculations of the finite density Thirring model. *Physical Review D*, 95(1):014502, 1 2017.

[35] Sumati Surya. The causal set approach to quantum gravity. *Living Reviews in Relativity*, 22(1):5, 12 2019.

[36] Neil Turok. On Quantum Tunneling in Real Time. *New Journal of Physics*, 16(6):063006, 12 2014.

[37] Aleksey Cherman and Mithat Unsal. Real-Time Feynman Path Integral Realization of Instantons. *arXiv: 1408.0012*.

[38] Yuya Tanizaki and Takayuki Koike. Real-time Feynman path integral with Picard–Lefschetz theory and its applications to quantum tunneling. *Annals of Physics*, 351:250–274, 6 2014.

[39] V.P. Frolov and G.A. Vilkovisky. Spherically symmetric collapse in quantum gravity. *Physics Letters B*, 106(4):307–313, 11 1981.

[40] V. P. Frolov, M. A. Markov, and V. F. Mukhanov. Through a black hole into a new universe? *Physics Letters B*, 216(3-4):272–276, 1 1989.

[41] Claude Barrabès and Valeri P. Frolov. How many new worlds are inside a black hole? *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 53(6):3215–3223, 3 1996.

[42] Valeri P. Frolov and Igor D. Novikov. *Black Hole Physics*, volume 96 of *Fundamental Theories of Physics*. Springer Netherlands, Dordrecht, 1998.

[43] Alexander Vilenkin. Creation of universes from nothing. *Physics Letters B*, 117(1-2):25–28, 11 1982.

[44] J. B. Hartle and S. W. Hawking. Wave function of the Universe. *Physical Review D*, 28(12):2960–2975, 12 1983.

[45] J. J. Halliwell. Introductory Lectures on Quantum Cosmology. In Piran T. Coleman S., Hartle J.B. and Weinberg S., editors, *Proceedings of the 1990 Jerusalem Winter School on Quantum Cosmology and Baby Universes*. World Scientific, Singapore, 1991.

[46] Martin Bojowald. Absence of a Singularity in Loop Quantum Cosmology. *Physical Review Letters*, 86(23):5227, 6 2001.

[47] Leonardo Modesto. Disappearance of the black hole singularity in loop quantum gravity. *Physical Review D*, 70(12):124009, 12 2004.

[48] Abhay Ashtekar and Martin Bojowald. Black hole evaporation: a paradigm. *Classical and Quantum Gravity*, 22(16):3349, 8 2005.

[49] Sean A. Hayward. Formation and Evaporation of Nonsingular Black Holes. *Physical Review Letters*, 96(3):031103, 1 2006.

[50] S. Hossenfelder and L. Smolin. Conservative solutions to the black hole information problem. *Physical Review D*, 81(6):064009, 3 2010.

[51] Hal M. Haggard and Carlo Rovelli. Quantum-gravity effects outside the horizon spark black to white hole tunneling. *Physical Review D*, 92(10):104020, 11 2015.

[52] Carlos Barceló, Raúl Carballo-Rubio, Luis J. Garay, and Gil Jannes. The lifetime problem of evaporating black holes: mutiny or resignation. *Classical and Quantum Gravity*, 32(3), 9 2014.

[53] Eugenio Bianchi, Marios Christodoulou, Fabio D'Ambrosio, Hal M Haggard, and Carlo Rovelli. White holes as remnants: a surprising scenario for the end of a black hole. *Classical and Quantum Gravity*, 35(22):225003, 10 2018.

[54] Fabio D'Ambrosio, Marios Christodoulou, Pierre Martin-Dussaud, Carlo Rovelli, and Farshid Soltani. End of a black hole's evaporation. *Physical Review D*, 103(10):106014, 5 2021.

[55] Daniele Oriti, Lorenzo Sindoni, and Edward Wilson-Ewing. Bouncing cosmologies from quantum gravity condensates. *Classical and Quantum Gravity*, 34(4):04LT01, 1 2017.

[56] James B. Hartle. Simplical minisuperspace. III. Integration contours in a five-simplex model. *Journal of Mathematical Physics*, 30(2):452–460, 2 1989.

[57] Li Xin Li, Jian Mei Xu, and Liao Liu. Complex geometry, quantum tunneling, and time machines. *Physical Review D*, 48(10):4735–4737, 11 1993.

[58] Steffen Gielen and Neil Turok. Perfect Quantum Cosmological Bounce. *Physical Review Letters*, 117(2), 10 2015.

[59] Steffen Gielen and Neil Turok. Quantum propagation across cosmological singularities. *Physical Review D*, 95(10), 12 2016.

[60] Job Feldbrugge, Jean-Luc Lehners, and Neil Turok. Lorentzian Quantum Cosmology. *Physical Review D*, 95(10), 3 2017.

[61] J. Diaz Dorronsoro, J. J. Halliwell, J. B. Hartle, T. Hertog, and O. Janssen. Real no-boundary wave function in Lorentzian quantum cosmology. *Physical Review D*, 96(4):043505, 8 2017.

[62] Sebastian F. Bramberger, Thomas Hertog, Jean-Luc Lehners, and Yannick Vreys. Quantum Transitions Through Cosmological Singularities. *Journal of Cosmology and Astroparticle Physics*, 2017(7), 1 2017.

[63] J Louko and P A Tuckey. Regge calculus in anisotropic quantum cosmology. *Classical and Quantum Gravity*, 9(1):41, 1 1992.

[64] Danny Birmingham. Lens spaces in the Regge calculus approach to quantum cosmology. *Physical Review D*, 52(10):5760, 11 1995.

[65] Danny Birmingham. A Closed Contour of Integration in Regge Calculus. *General Relativity and Gravitation 1998 30:1*, 30(1):83–103, 1998.

[66] Yasuhiko Furihata. No-boundary wave function for a simplicial anisotropic universe. *Physical Review D*, 53(12):6875, 6 1996.

[67] Cristóvão Correia da Silva and Ruth M Williams. Simplicial minisuperspace models in the presence of a scalar field. *Classical and Quantum Gravity*, 16(7):2197, 7 1999.

[68] Cristóvão Correia da Silva and Ruth M Williams. Anisotropic simplicial minisuperspace model in the presence of a scalar field. *Classical and Quantum Gravity*, 16(8):2681, 8 1999.

[69] Cristóvão Correia da Silva and Ruth M Williams. Simplicial minisuperspace models in the presence of a massive scalar field with arbitrary scalar coupling eta R phiˆ2. *Classical and Quantum Gravity*, 17(8):1827, 4 2000.

[70] Cristovao Correia da Silva and Ruth M. Williams. Wormholes in Simplicial Minisuperspace. *arXiv: gr-qc/0004035*.

[71] D. Sorkin, Rafael. Lorentzian angles and trigonometry including lightlike vectors. *arXiv: 1908.10022*.

[72] James B. Hartle. Simplicial minisuperspace I. General discussion. *Journal of Mathematical Physics*, 26(4):804–814, 4 1985.

[73] Bianca Dittrich and Sebastian Steinhaus. Path integral measure and triangulation independence in discrete gravity. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 85(4):044032–undefined, 10 2011.

[74] Bianca Dittrich, Wojciech Kaminski, and Sebastian Steinhaus. Discretization independence implies non-locality in 4D discrete quantum gravity. *Classical and Quantum Gravity*, 31(24):245009, 4 2014.

[75] J. B. Hartle. Unruly topologies in two-dimensional quantum gravity. *Classical and Quantum Gravity*, 2(5):707–720, 9 1985.

[76] D. Sorkin, Rafael. *Development of simplectic methods for the metrical and electromagnetic fields*. PhD thesis, California Institute of Technology, 1974.

[77] Jorma Louko and Rafael D. Sorkin. Complex actions in two-dimensional topology change. *Classical and Quantum Gravity*, 14(1):179–204, 11 1995.

[78] Casey E. Berger, Lukas Rammelmüller, Andrew C. Loheac, Florian Ehmann, Jens Braun, and Joaquín E. Drut. Complex Langevin and other approaches to the sign problem in quantum many-body physics. *Physics Reports*, 892:1–54, 7 2019.

[79] Christof Gattringer and Kurt Langfeld. Approaches to the sign problem in lattice field theory. *International Journal of Modern Physics A*, 31(22):1643007–, 3 2016.

[80] Masafumi Fukuma and Naoya Umeda. Parallel tempering algorithm for integration over Lefschetz thimbles. *Progress of Theoretical and Experimental Physics*, 2017(7), 3 2017.

[81] Andrei Alexandru, Gokce Basar, Paulo F. Bedaque, and Neill C. Warrington. Tempered transitions between thimbles. *Physical Review D*, 96(3), 3 2017.

[82] Herbert W. Hamber and Ruth M. Williams. Two-dimensional simplical quantum gravity. *Nuclear Physics, Section B*, 267(2):482–496, 4 1986.

[83] Gareth O. Roberts and Jeffrey S. Rosenthal. Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.

[84] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, 59(1):65–98, 2 2017.

[85] Masafumi Fukuma and Nobuyuki Matsumoto. Worldvolume approach to the tempered Lefschetz thimble method. *Progress of Theoretical and Experimental Physics*, 2021(2):23–31, 2 2021.

[86] Masafumi Fukuma, Nobuyuki Matsumoto, and Yusuke Namekawa. Statistical analysis method for the worldvolume hybrid Monte Carlo algorithm. *arXiv:2107.06858*, 7 2021.

[87] Scott Lawrence and Yukari Yamauchi. Normalizing flows and the real-time sign problem. *Physical Review D*, 103(11):114509, 6 2021.

[88] Jan-Lukas Wynen, Evan Berkowitz, Stefan Krieg, Thomas Luu, and Johann Ostmeyer. Machine learning to alleviate Hubbard-model sign problems. *Physical Review B*, 103(12):125153, 3 2021.

[89] Daniele Malafarina. Classical Collapse to Black Holes and Quantum Bounces: A Review. *Universe*, 3(2):48, 5 2017.

[90] Viatcheslav Mukhanov. *Physical foundations of cosmology*. Cambridge University Press, Cambridge, 2005.

[91] Eugenio Bianchi and Pierre Martin-Dussaud. Causal structure in spin-foams. *arxiv: 2109.00986*.

[92] Fay Dowker and Jeremy Butterfield. Recovering General Relativity from a Planck scale discrete theory of quantum gravity. *arXiv: 2106.01297*.

[93] Seth K. Asante, Bianca Dittrich, and José Padua-Argüelles. Complex actions and causality violations: Applications to Lorentzian quantum cosmology. *Classical and Quantum Gravity*, 40(10):105005, 4 2023.

[94] D. Sorkin, Rafael. Time-evolution problem in regge calculus. *Physical Review D*, 12(2):385–396, 7 1975.

[95] Jan Ambjorn, Jakub Gizbert-Studnicki, Andrzej Görlich, Jerzy Jurkiewicz, and Renate Loll. Renormalization in Quantum Theories of Geometry. *Frontiers in Physics*, 8:247, 7 2020.

[96] Sebastian Steinhaus. Coarse Graining Spin Foam Quantum Gravity—A Review. *Frontiers in Physics*, 8:295, 8 2020.

[97] Istvan Montvay and Gernot Münster. *Quantum Fields on a Lattice*. Cambridge University Press, Cambridge, 1994.

[98] David J.E. Callaway. Triviality pursuit: Can elementary scalar particles exist? *Physics Reports*, 167(5):241–320, 1988.

[99] Jan Ambjorn. Quantization of Geometry. *arXiv: hep-th/9411179*.

[100] Tullio Regge and Ruth M. Williams. Discrete structures in gravity. *Journal of Mathematical Physics*, 41(6):3964–3984, 12 2000.

[101] J. Ambjørn and R. Loll. Non-perturbative Lorentzian quantum gravity, causality and topology change. *Nuclear Physics B*, 536(1-2):407–434, 12 1998.

[102] Renate Loll. A Discrete History of the Lorentzian Path Integral. In Domenico J. W. Giulini, Claus Kiefer, and Claus Lämmerzahl, editors, *Quantum Gravity: From Theory to Experimental Search*, pages 137–171. Springer, Berlin, 2003.

[103] Lee Smolin. Quantum gravity on a lattice. *Nuclear Physics B*, 148(3-4):333–372, 2 1979.

[104] James B. Hartle. Simplicial minisuperspace. II. Some classical solutions on simple triangulations. *Journal of Mathematical Physics*, 27(1):287–295, 1 1986.

[105] Jonathan J. Halliwell and Jorma Louko. Steepest-descent contours in the path-integral approach to quantum cosmology. I. The de sitter minisuperspace model. *Physical Review D*, 39(8):2206–2215, 4 1989.

[106] Jonathan J. Halliwell and Jorma Louko. Steepest-descent contours in the path-integral approach to quantum cosmology. II. Microsuperspace. *Physical Review D*, 40(6):1868, 9 1989.

[107] Jonathan J. Halliwell and Jorma Louko. Steepest-descent contours in the path-integral approach to quantum cosmology. III. A general method with applications to anisotropic minisuperspace models. *Physical Review D*, 42(12):3997, 12 1990.

[108] Jonathan J. Halliwell and Robert C. Myers. Multiple-sphere configurations in the path-integral representation of the wave function of the Universe. *Physical Review D*, 40(12):4011, 12 1989.

[109] Jonathan J. Halliwell and James B. Hartle. Integration contours for the no-boundary wave function of the universe. *Physical Review D*, 41(6):1815, 3 1990.

[110] Kristin Schleich. Semiclassical wave function of the Universe at small three-geometries. *Physical Review D*, 32(8):1889, 10 1985.

[111] Kristin Schleich. Conformal rotation in Bianchi type-I quantum cosmology. *Physical Review D*, 39(8):2192, 4 1989.

[112] Edward Farhi. The wave function of the universe and the square root of minus one. *Physics Letters B*, 219(4):403–407, 3 1989.

[113] James M. Cline. Does the wormhole mechanism for vanishing cosmological constant work in lorentzian gravity? *Physics Letters B*, 224(1-2):53–57, 6 1989.

[114] Wai Mo Suen and Kenneth Young. Wave function of the Universe as a leaking system. *Physical Review D*, 39(8):2201, 4 1989.

[115] J. David Brown and Erik A. Martinez. Lorentzian path integral for minisuperspace cosmology. *Physical Review D*, 42(6):1931, 9 1990.

[116] Rafael D. Sorkin. Is the spacetime metric Euclidean rather than Lorentzian? In Arundhati Dasgupta, editor, *Recent Research in Quantum Gravity*, pages 115–135. Nova Science Publishers, New York, 2013.

[117] Job Feldbrugge, Jean-Luc Lehners, and Neil Turok. No smooth beginning for spacetime. *Physical Review Letters*, 119(17), 4 2017.

[118] Job Feldbrugge, Jean-Luc Lehners, and Neil Turok. No Rescue for the No Boundary Proposal. *Physical Review D*, 97(2), 8 2017.

[119] J. Diaz Dorronsoro, J. J. Halliwell, J. B. Hartle, T. Hertog, O. Janssen, and Y. Vreys. Damped Perturbations in the No-Boundary State. *Physical Review Letters*, 121(8):081302, 8 2018.

[120] Job Feldbrugge, Jean Luc Lehners, and Neil Turok. Inconsistencies of the new no-boundary proposal. *Universe*, 4(10), 10 2018.

[121] Alexander Vilenkin and Masaki Yamada. Tunneling wave function of the universe. *Physical Review D*, 98(6):066003, 9 2018.

[122] Alice Di Tucci and Jean Luc Lehners. Unstable no-boundary fluctuations from sums over regular metrics. *Physical Review D*, 98(10):103506, 11 2018.

[123] Alice Di Tucci, Job Feldbrugge, Jean-Luc Lehners, and Neil Turok. Quantum Incompleteness of Inflation. *Physical Review D*, 100(6), 6 2019.

[124] Oliver Janssen, Jonathan J. Halliwell, and Thomas Hertog. The no-boundary proposal in biaxial Bianchi IX minisuperspace. *Physical Review D*, 99(12), 4 2019.

[125] Alice Di Tucci and Jean Luc Lehners. No-Boundary Proposal as a Path Integral with Robin Boundary Conditions. *Physical Review Letters*, 122(20):201302, 5 2019.

[126] Alice Di Tucci, Jean-Luc Lehners, and Laura Sberna. No-boundary prescriptions in Lorentzian quantum cosmology. *Physical Review D*, 100(12), 11 2019.

[127] Sebastian F. Bramberger, Alice Di Tucci, and Jean-Luc Lehners. Homogeneous Transitions during Inflation: a Description in Quantum Cosmology. *Physical Review D*, 101(6), 7 2019.

[128] Alice Di Tucci, Michal P. Heller, and Jean Luc Lehners. Lessons for quantum cosmology from anti-de Sitter black holes. *Physical Review D*, 102(8):086011, 10 2020.

[129] Gaurav Narain and Hai-Qing Zhang. Lorentzian quantum cosmology in novel Gauss-Bonnet gravity from Picard-Lefschetz methods. *arXiv: 2006.02298*.

[130] Karthik Rajeev, Vikramaditya Mondal, and Sumanta Chakraborty. No-boundary Wave Function, Wheeler-DeWitt Equation and Path Integral Analysis of the Bouncing 'Quantum' Cosmology. *Physical Review D*, 103(10), 1 2021.

[131] Karthik Rajeev, Vikramaditya Mondal, and Sumanta Chakraborty. Bouncing with shear: Implications from quantum cosmology. *Journal of Cosmology and Astroparticle Physics*, 2022(1), 9 2021.

[132] Jean-Luc Lehners. The Wave Function of Simple Universes, Analytically Continued From Negative to Positive Potentials. *Physical Review D*, 104(6), 5 2021.

[133] Gaurav Narain. On Gauss-Bonnet gravity and boundary conditions in Lorentzian path-integral quantization. *Journal of High Energy Physics 2021 2021:5*, 2021(5):1–35, 5 2021.

[134] Jean Luc Lehners. Allowable complex metrics in minisuperspace quantum cosmology. *Physical Review D*, 105(2):026022, 1 2022.

[135] Caroline Jonas, Jean-Luc Lehners, and Vincent Meyer. Revisiting the no-boundary proposal with a scalar field. *Physical Review D*, 105(4):043529, 2 2022.

[136] Hiroki Matsui. Lorentzian path integral for quantum tunneling and WKB approximation for wave-function. *The European Physical Journal C 2022 82:5*, 82(5):1–16, 5 2022.

[137] Gaurav Narain. Surprises in Lorentzian path-integral of Gauss-Bonnet gravity. *Journal of High Energy Physics 2022 2022:4*, 2022(4):1–27, 4 2022.

[138] Caroline Jonas, Jean-Luc Lehners, and Jerome Quintin. Uses of complex metrics in cosmology. *Journal of High Energy Physics 2022 2022:8*, 2022(8):1–39, 8 2022.

[139] Jonathan J. Halliwell. Derivation of the Wheeler-DeWitt equation from a path integral for minisuperspace models. *Physical Review D*, 38(8):2468, 10 1988.

[140] James W. York. Role of conformal three-geometry in the dynamics of gravitation. *Physical Review Letters*, 28(16):1082–1085, 4 1972.

[141] G. W. Gibbons and S. W. Hawking. Action integrals and partition functions in quantum gravity. *Physical Review D*, 15(10):2752–2756, 5 1977.

[142] Alexander Vilenkin. Quantum creation of universes. *Physical Review D*, 30(2):509–511, 7 1984.

[143] A. D. Linde. Quantum creation of the inflationary universe. *Lettere al Nuovo Cimento*, 39(17):401–405, 4 1984.

[144] Aurélien Barrau, Killian Martineau, and Flora Moulin. Seeing through the cosmological bounce: Footprints of the contracting phase and luminosity distance in bouncing models. *Physical Review D*, 96(12):123520, 12 2017.

[145] Genki Fujisawa, Jun Nishimura, Katsuta Sakai, and Atis Yosprakob. Backpropagating Hybrid Monte Carlo algorithm for fast Lefschetz thimble calculations. *Journal of High Energy Physics*, 2022(4):1–27, 4 2022.

[146] Christopher Rackauckas, Anshul Singhvi, Yingbo Ma, Chris de Graaf, Lilith Orion Hafner, Michael Hatherly, Scott P. Jones, dextorious, Arno Strouwen, Colin Caine, Elliot Saba, Julia TagBot, Kvaz1r, Max G, Sam Isaacson, Sheehan Olver, The Gitter Badger, c123w, and Maja Gwóźdź. SciML/DifferentialEquations.jl: v7.3.0. 8 2022.

[147] Eric Delabaere and C. J. Howls. Global asymptotics for multiple integrals with boundaries. *Duke Mathematical Journal*, 112(2):199–264, 4 2002.

[148] Alexander Vilenkin. Quantum cosmology and the initial state of the Universe. *Physical Review D*, 37(4):888, 2 1988.

[149] Georgios Fanaras and Alexander Vilenkin. The tunneling wavefunction in Kantowski-Sachs quantum cosmology. *Journal of Cosmology and Astroparticle Physics*, 2022(08):069, 8 2022.

[150] Robert Brandenberger and Patrick Peter. Bouncing Cosmologies: Progress and Problems. *Foundations of Physics 2017 47:6*, 47(6):797–850, 2 2017.

[151] Yoshiyasu Ito, Daisuke Kadoh, and Yuki Sato. Tensor network approach to 2D Lorentzian quantum Regge calculus. *Physical Review D*, 106(10):106004, 11 2022.

[152] Latham Boyle and Neil Turok. Two-Sheeted Universe, Analyticity and the Arrow of Time. *arXiv:2109.06204*.

[153] Neil Turok and Latham Boyle. Gravitational entropy and the flatness, homogeneity and isotropy puzzles. *arXiv:2201.07279*.

[154] Latham Boyle and Neil Turok. Thermodynamic solution of the homogeneity, isotropy and flatness puzzles (and a clue to the cosmological constant). *arXiv:2210.01142*.

[155] Mari Carmen Bauls and Krzysztof Cichy. Review on novel methods for lattice gauge theories. *Reports on Progress in Physics*, 83(2):024401, 1 2020.

[156] Abhay Ashtekar, Tomasz Pawlowski, and Parampreet Singh. Quantum nature of the big bang: An analytical and numerical investigation. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 73(12):124038, 2006.

[157] Abhay Ashtekar. Singularity Resolution in Loop Quantum Cosmology: A Brief Overview. *Journal of Physics: Conference Series*, 189:012003, 2009.

[158] Johanna N. Borissova and Astrid Eichhorn. Towards black-hole singularity-resolution in the Lorentzian gravitational path integral. *Universe*, 7(3), 2020.

[159] Gary T. Horowitz and Juan Maldacena. The black hole final state. *Journal of High Energy Physics*, 8(2):197–211, 2003.

[160] Malcolm J. Perry. Future Boundaries and the Black Hole Information Paradox. *arXiv:2108.05744*.

[161] Eric Greenwood and Dejan Stojkovic. Quantum gravitational collapse: non-singularity and non-locality. *Journal of High Energy Physics*, 2008(6):042, 6 2008.

[162] John E. Wang, Eric Greenwood, and Dejan Stojkovic. Schrodinger formalism, black hole horizons, and singularity behavior. *Physical Review D*, 80(12):124027, 12 2009.

[163] Anshul Saini and Dejan Stojkovic. Nonlocal (but also nonsingular) physics at the last stages of gravitational collapse. *Physical Review D*, 89(4):044003, 2 2014.

[164] Robert Wald. *General Relativity*. University of Chicago Press, Chicago, 1984.

[165] S. Jordan and R. Loll. Causal Dynamical Triangulations without preferred foliation. *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, 724(1-3):155–159, 7 2013.

[166] S. Jordan and R. Loll. De Sitter universe from causal dynamical triangulations without preferred foliation. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 88(4):044055, 8 2013.

[167] R. Loll and W. Westra. Sum over topologies and double-scaling limit in 2D Lorentzian quantum gravity. *Classical and Quantum Gravity*, 23(2):465–471, 1 2006.

[168] Rafael D. Sorkin. Consequences of spacetime topology. In A. Coley, F. Cooperstock, and B. Tupper, editors, *Proceedings of the Third Canadian Conference on General Relativity and Relativistic Astrophysics*, pages 137–163. World Scientific, 1990.

[169] Donald Marolf and Henry Maxfield. Observations of Hawking radiation: the Page curve and baby universes. *Journal of High Energy Physics 2021 2021:4*, 2021(4):1–79, 4 2021.

[170] Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. *Reviews of Modern Physics*, 93(3), 6 2020.

[171] Carlo Rovelli. The Subtle Unphysical Hypothesis of the Firewall Theorem. *Entropy*, 21(9), 2019.

[172] Julian Schwinger. Brownian motion of a quantum oscillator. *Journal of Mathematical Physics*, 2(3):407–432, 5 1961.

[173] L V Keldysh. Diagram technique for nonequilibrium processes. *J. Exptl. Theoret. Phys. (U.S.S.R.)*, 20(4):1515–1527, 1965.

[174] R.P Feynman and F.L Vernon. The theory of a general quantum system interacting with a linear dissipative system. *Annals of Physics*, 24:118–173, 10 1963.

[175] Edward Witten. 2 + 1 dimensional gravity as an exactly soluble system. *Nuclear Physics B*, 311(1):46–78, 12 1988.

[176] Edward Witten. Three-Dimensional Gravity Revisited. *arXiv: 0706.3359*.

[177] S. W. Hawking. The quantum state of the universe. *Nuclear Physics B*, 239(1):257–276, 6 1984.

[178] Stephen Hawking and Roger Penrose. *The nature of space and time*. Princeton University Press, 2010.

[179] Pisin Chen, Yao Chieh Hu, and Dong Han Yeom. Fuzzy Euclidean wormholes in de Sitter space. *Journal of Cosmology and Astroparticle Physics*, 2017(07):001, 7 2017.

[180] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction To Quantum Field Theory*. Westview Press, 1995.

[181] Alfio Bonanno, Astrid Eichhorn, Holger Gies, Jan M. Pawlowski, Roberto Percacci, Martin Reuter, Frank Saueressig, and Gian Paolo Vacca. Critical reflections on asymptotically safe gravity. *Frontiers in Physics*, 8:269, 8 2020.

[182] Matt Visser. How to Wick rotate generic curved spacetime. *arXiv.1702.05572*.

[183] John F. Donoghue. A Critique of the Asymptotic Safety Program. *Frontiers in Physics*, 8:56, 3 2020.

[184] Pisin Chen, Misao Sasaki, Dong-han Yeom, and Junggi Yoon. Solving information loss paradox via Euclidean path integral. *arXiv:2111.01005*.

[185] Don N. Page. Quantum Cosmology. In G. W. Gibbons, E. P. S. Shellard, and S. J. Rankin, editors, *The Future of Theoretical Physics and Cosmology: Celebrating Stephen Hawking's Contributions to Physics*, chapter 35, pages 621–648. Cambridge University Press, 2003.

[186] Daniele Oriti, editor. *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*. Cambridge University Press, 2009.

[187] James B. Hartle. Spacetime Quantum Mechanics and the Quantum Mechanics of Spacetime. In B Julia and J Zinn-Justin, editors, *Gravitation and Quantizations: Proceedings of the 1992 Les Houches Summer School*. North Holland, Amsterdam, 1995.

[188] Mario Castagnino, Sebastian Fortin, Roberto Laura, and Daniel Sudarsky. Interpretations of Quantum Theory in the Light of Modern Cosmology. *Foundations of Physics*, 47(11):1387–1422, 11 2017.

[189] Sean M. Carroll. Why Boltzmann Brains Are Bad. *Current Controversies in Philosophy of Science*, pages 7–20, 2017.

[190] Rafael D. Sorkin. Impossible Measurements on Quantum Fields. In B.L. Hu; T.A. Jacobson, editor, *Directions in General Relativity, Proceedings of the 1993 International Symposium, Maryland: Papers in Honor of Dieter Brill*, volume 2, pages 293–305. Cambridge University Press, Cambridge, 1993.

[191] L. Borsten, I. Jubb, and G. Kells. Impossible measurements revisited. *Physical Review D*, 104(2):025012, 7 2021.

[192] I. Jubb. Causal state updates in real scalar quantum field theory. *Physical Review D*, 105(2):025003, 1 2022.

[193] E P Wigner. Remarks on the Mind-Body Question. In I.J. Good, editor, *The Scientist Speculates*. Heinemann, London, 1961.

[194] Časlav Brukner. On the Quantum Measurement Problem. In Reinhold Bertlmann and Anton Zeilinger, editors, *Quantum [Un]Speakables II*, pages 95–117. Springer, Basel, 2017.

[195] Časlav Brukner. A No-Go Theorem for Observer-Independent Facts. *Entropy 2018, Vol. 20, Page 350*, 20(5):350, 5 2018.

[196] Daniela Frauchiger and Renato Renner. Quantum theory cannot consistently describe the use of itself. *Nature Communications*, 9(1):1–10, 9 2018.

[197] Matthew Pusey. Is QBism 80% complete, or 20%? *Talk at Rotman Institute of Philosophy (https://www.youtube.com/watch?v=_9Rs61l8MyY)*, 2016.

[198] Kok Wei Bong, Aníbal Utreras-Alarcón, Farzad Ghafari, Yeong Cherng Liang, Nora Tischler, Eric G. Cavalcanti, Geoff J. Pryde, and Howard M. Wiseman. A strong no-go theorem on the Wigner's friend paradox. *Nature Physics*, 16(12):1199–1205, 12 2020.

[199] Marek Zukowski and Marcin Markiewicz. Physics and Metaphysics of Wigner's Friends: Even Performed Premeasurements Have No Results. *Physical Review Letters*, 126(13):130402, 4 2021.

[200] Don N. Page. Sensible Quantum Mechanics: Are Only Perceptions Probabilistic? *arXiv: quant-ph/9506010*.

[201] Don N. Page. Sensible Quantum Mechanics: Are Probabilities only in the Mind? *International Journal of Modern Physics D*, 5(6):583–596, 7 1995.

[202] Don N. Page. Mindless Sensationalism: A Quantum Framework for Consciousness. In Quentin Smith and Alexsandar Jokic, editors, *Consciousness: new philosophical perspectives*, chapter 17, pages 468–506. Oxford University Press, Oxford, 2003.

[203] Don N. Page. Consciousness and the Quantum. *arXiv: 1102.5339*.

[204] Don N. Page. Born's Rule Is Insufficient in a Large Universe. *arXiv: 1003.2419*.

[205] Fay Dowker and Adrian Kent. On the Consistent Histories Approach to Quantum Mechanics. *Journal of Statistical Physics*, 82(5-6):1575–1646, 12 1996.

[206] Robert Oeckl. A local and operational framework for the foundations of physics. *Advances in Theoretical and Mathematical Physics*, 23(2):437–592, 2019.

[207] Carlo Rovelli. The Relational Interpretation. In Olival Freire, editor, *The Oxford Handbook of the History of Quantum Interpretations*, pages 1055–1071. Oxford University Press, Oxford, 2022.

[208] D. J. Chalmers. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–219, 3 1995.

[209] Daniel C. Dennett. Facing up to the hard question of consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755), 9 2018.

[210] Lawrence Shapiro and Spaulding Shannon. Embodied Cognition. In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy*. 2021.

[211] Peter Godfrey-Smith. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Macmillan, New York, 2016.

[212] David J. Chalmers and Kelvin J. McQueen. Consciousness and the Collapse of the Wave Function. In Shan Gao, editor, *Consciousness and Quantum Mechanics*, pages 11–63. Oxford University Press, Oxford, 2022.

[213] Kobi Kremnizer and André Ranchin. Integrated Information-Induced Quantum Collapse. *Foundations of Physics*, 45(8):889–899, 8 2015.

[214] Elias Okon and Miguel Ángel Sebastián. A consciousness-based quantum objective collapse model. *Synthese*, 197(9):3947–3967, 9 2020.

[215] Elias Okon and Miguel Ángel Sebastián. The Subjective-Objective Collapse Model Virtues and Challenges. In Shan Gao, editor, *Consciousness and Quantum Mechanics*, pages 64–82. Oxford University Press, Oxford, 2022.

[216] Adrian Kent. Collapse and Measures of Consciousness. *Foundations of Physics 2021 51:3*, 51(3):1–14, 5 2021.

[217] John B. DeBrota, Christopher A. Fuchs, and Rüdiger Schack. Respecting One's Fellow: QBism's Analysis of Wigner's Friend. *Foundations of Physics*, 50(12):1859–1874, 12 2020.

[218] Veronika Baumann and Časlav Brukner. Wigner's Friend as a Rational Agent. In Meir Hemmo and Orly Shenker, editors, *Quantum, Probability, Logic*, pages 91–99. Springer, Cham, 2020.

[219] Elias Okon. On the objectivity of measurement outcomes. *The British Journal for the Philosophy of Science*, (forthcoming).

[220] Elias Okon. Reassessing the strength of a class of Wigner's friend no-go theorems. *arXiv: 2204.12015*.

[221] Christopher A. Fuchs and Blake C. Stacey. QBism: Quantum Theory as a Hero's Handbook. *Proceedings of the International School of Physics "Enrico Fermi"*, 197:133–202, 12 2016.

[222] Christopher A. Fuchs. Notwithstanding Bohr, the Reasons for QBism. *Mind and Matter*, 15(2):245–300, 5 2017.

[223] Blake C. Stacey. Ideas Abandoned en Route to QBism. *arXiv: 1911.07386*.

[224] Lee Smolin. *Einstein's Unfinished Revolution*. Penguine, London, 2019.

[225] Rafael D. Sorkin. Forks in the Road, on the Way to Quantum Gravity. *International Journal of Theoretical Physics*, 36(12):2759–2781, 1997.

[226] Stefan Teufel and Detlef Dürr. *Bohmian mechanics: The physics and mathematics of quantum theory*. Springer, Berlin, 2009.

[227] Angelo Bassi, Kinjalk Lochan, Seema Satin, Tejinder P. Singh, and Hendrik Ulbricht. Models of wave-function collapse, underlying theories, and experimental tests. *Reviews of Modern Physics*, 85(2):471–527, 4 2013.

[228] David Wallace. On the Plurality of Quantum Theories: Quantum Theory as a Framework, and its Implications for the Quantum Measurement Problem. In Steven French and Juha Saatsi, editors, *Scientific Realism and the Quantum*, pages 78–102. Oxford University Press, Oxford, 2 2020.

[229] David Wallace. The sky is blue, and other reasons quantum mechanics is not underdetermined by evidence. *arXiv: 2205.00568*.

[230] Hugh Everett. "Relative State" Formulation of Quantum Mechanics. *Reviews of Modern Physics*, 29(3):454, 7 1957.

[231] Claus Kiefer. *Quantum Gravity*. Oxford University Press, Oxford, 3rd edition, 2012.

[232] Jonathan Barrett, Adrian Kent, Simon Saunders, and David Wallace, editors. *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, Oxford, 2010.

[233] David Wallace. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford University Press, Oxford, 2012.

[234] Murray Gell-Mann and James B. Hartle. Adaptive Coarse Graining, Environment, Strong Decoherence, and Quasiclassical Realms. *Physical Review A - Atomic, Molecular, and Optical Physics*, 89(5), 12 2013.

[235] Federico Laudisa and Carlo Rovelli. Relational Quantum Mechanics. In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy*. 2021.

[236] Carlo Rovelli and Francesca Vidotto. *Covariant Loop Quantum Gravity*. Cambridge University Press, Cambridge, 2014.

[237] James B. Hartle. Quantum pasts and the utility of history. *Physica Scripta T*, 76(T76):67–77, 1 1998.

[238] Ricardo Muciño, Elias Okon, and Daniel Sudarsky. Assessing relational quantum mechanics. *Synthese*, 200(5):1–26, 10 2022.

[239] Časlav Brukner. Qubits are not observers – a no-go theorem. *arXiv:2107.03513*.

[240] Emily Adlam and Carlo Rovelli. Information is Physical: Cross-Perspective Links in Relational Quantum Mechanics. *arXiv:2203.13342*.

[241] Hervé Zwirn. Is QBism a Possible Solution to the Conceptual Problems of Quantum Mechanics? In Olival Freire, editor, *The Oxford Handbook of the History of Quantum Interpretations*. Oxford University Press, Oxford, 2022.

[242] Humberto R. Maturana and Francisco J. Varela. *Autopoiesis and Cognition*. Boston Studies in the Philosophy and History of Science. Springer Netherlands, Dordrecht, 1980.

[243] Fritjof Capra and Pier Luigi Luisi. *The systems view of life: A unifying vision*. Cambridge University Press, Cambridge, 2012.

[244] Gennaro Auletta. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford University Press, Oxford, 2011.

[245] William Bialek. *Biophysics: Searching for principles*. Princeton University Press, Princeton, 2012.

[246] Adrian Kent. Night thoughts of a quantum physicist. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 358(1765):75–87, 2000.

[247] D. L. Wiltshire. An introduction to quantum cosmology. In Visvanathan N. Robson B. and Woolcock W.S., editors, *Cosmology: the Physics of the Universe*, pages 473–531. World Scientific, Singapore, 1996.

[248] Alexander Vilenkin. Approaches to Quantum Cosmology. *Physical Review D*, 50(4):2581–2594, 3 1994.

[249] Jonathan J. Halliwell, James B. Hartle, and Thomas Hertog. What is the No-Boundary Wave Function of the Universe? *Physical Review D*, 99(4):043526, 2 2019.

[250] Jean-Luc Lehners. Review of the No-Boundary Wave Function. *arXiv: 2303.08802*.

[251] Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz. Time Symmetry in the Quantum Process of Measurement. *Physical Review*, 134(6B):B1410, 6 1964.

[252] Robert B. Griffiths. Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics 1986 36:1*, 36(1):219–272, 7 1984.

[253] Murray Gell-Mann and James B. Hartle. Time Symmetry and Asymmetry in Quantum Mechanics and Quantum Cosmology. In J. Halliwell, J. Perez-Mercader, and W. Zurek, editors, *Physical origins of time asymmetry*, pages 311–345. Cambridge University Press, Cambridge, 1994.

[254] Murray Gell-Mann and James B. Hartle. Quantum Mechanics in the Light of Quantum Cosmology. In Wojciech H. Zurek, editor, *Complexity, Entropy, and the Physics of Information*, pages 425–458. CRC Press, Boca Raton, 1990.

[255] Roland Omnès. Logical reformulation of quantum mechanics. I. Foundations. *Journal of Statistical Physics 1988 53:3*, 53(3):893–932, 11 1988.

[256] David Craig and James B. Hartle. Generalized quantum theory of recollapsing homogeneous cosmologies. *Physical Review D*, 69(12):123525, 6 2004.

[257] David A. Craig and Parampreet Singh. Consistent probabilities in Wheeler-DeWitt quantum cosmology. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 82(12):123526, 12 2010.

[258] J. J. Halliwell. Probabilities in quantum cosmological models: A decoherent histories analysis using a complex potential. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 80(12):124032, 12 2009.

[259] Charis Anastopoulos and Ntina Savvidou. Minisuperspace models in histories theory. *Classical and Quantum Gravity*, 22(9):1841, 4 2005.

[260] Richard P Feynman. *The Character of Physical Law*. MIT Press, Cambridge, Massachusetts, 1967.

[261] David Z. Albert. *Time and chance*. Harvard University Press, Cambridge, Massachusetts, 2000.

[262] Meir Hemmo and Orly Shenker. The Second Law of Thermodynamics and the Psychological Arrow of Time. *The British Journal for the Philosophy of Science*, 8 2019.

[263] Carlo Rovelli. Why do we remember the past and not the future? The 'time oriented coarse graining' hypothesis. *arXiv: 1407.3384*.

[264] Carlo Rovelli. Is time's arrow perspectival? In *The Philosophy of Cosmology*, pages 285–296. Cambridge University Press, 1 2017.

[265] R. P. Feynman. Mathematical formulation of the quantum theory of electromagnetic interaction. *Physical Review*, 80(3):440–457, 11 1950.

[266] Richard P. Feynman. An operator calculus having applications in quantum electrodynamics. *Physical Review*, 84(1):108–128, 1951.

[267] Kenneth G. Wilson. Confinement of quarks. *Physical Review D*, 10(8):2445–2459, 10 1974.

[268] Hagen Kleinert. *Gauge Fields in Condensed Matter*. World Scientific, Singapore, 1989.

[269] Roberto Fernández, Jürg Fröhlich, and Alan D. Sokal. *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory*. Springer-Verlag, Berlin, 1992.

[270] Bernd Bruegmann and E. Marinari. 4d Simplicial Quantum Gravity with a Non-Trivial Measure. *Physical Review Letters*, 70(13):1908–1911, 10 1992.

[271] Jan Ambjørn, Bergfinnur Durhuus, and Thordur Jonsson. *Quantum Geometry*. Cambridge University Press, Cambridge, 1997.

[272] Kevin Costello. *Renormalization and effective field theory*. American Mathematical Society, 2011.

[273] James P. Edwards and Christian Schubert. Quantum mechanical path integrals in the first quantised approach to quantum field theory. *arXiv:1912.10004*.

[274] Christof Gattringer and Thomas Kloiber. Spectroscopy in finite density lattice field theory: An exploratory study in the relativistic Bose gas. *Physics Letters B*, 720:210–214, 2013.

[275] Ydalia Delgado Mercado, Christof Gattringer, and Alexander Schmidt. Surface worm algorithm for abelian Gauge-Higgs systems on the lattice. *Computer Physics Communications*, 184(6):1535–1546, 6 2013.

[276] Christof Gattringer, Daniel Göschl, and Carlotta Marchis. Worldlines and worldsheets for non-abelian lattice field theories: Abelian color fluxes and Abelian color cycles. *EPJ Web of Conferences*, 175:11007, 3 2018.

[277] Carlotta Marchis and Christof Gattringer. Dual representation of lattice QCD with worldlines and worldsheets of Abelian color fluxes. *Physical Review D*, 97(3):034508, 2 2018.

[278] Carlo Rovelli. Why Gauge? *Foundations of Physics 2013 44:1*, 44(1):91–104, 1 2014.

[279] Steven Weinberg. *The Quantum Theory of Fields*. Cambridge University Press, Cambridge, 1995.

[280] Matthew Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Cambridge, 2013.