ABSTRACT

Title of Dissertation: QUANTIZATION OF CAUSAL DIAMONDS IN (2+1)-DIMENSIONAL GRAVITY

Rodrigo Andrade e Silva
Doctor of Philosophy, 2024

Dissertation Directed by: Professor Ted Jacobson
Department of Physics

We develop the non-perturbative reduced phase space quantization of *causal diamonds* in $(2+1)$-dimensional gravity with a nonpositive cosmological constant. The system is defined as the domain of dependence of a spacelike topological disk with fixed boundary metric. By solving the constraints in a constant-mean-curvature time gauge and removing all the spatial gauge redundancy, we find that the phase space is the cotangent bundle of $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. Classically, the states correspond to causal diamonds embedded in AdS_3 (or Mink_3 if $\Lambda = 0$), with fixed corner length, and whose Cauchy surfaces have the topology of a disc. Because the phase space does not have a natural linear structure, a generalization of the standard canonical (coordinate) quantization is required. As the configuration space is a homogeneous space for the $\mathit{Diff}^+(S^1)$ group, we apply Isham's group-theoretic quantization scheme. We propose a quantization based on (projective) unitary irreducible representations of the BMS_3 group. We find a class of suitable quantum theories labelled by a choice of a coadjoint orbit of the Virasoro group and an irreducible unitary representation of the corresponding little group. The most natural choice, justified by a Casimir matching principle, corresponds to a Hilbert space realized by wavefunctions on $\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$ valued in some unitary irreducible representation of $SL(2,\mathbb{R})$. A surprising result is that the twist of the diamond corner loop is quantized in terms of the ratio of the Planck length to the corner perimeter.

QUANTIZATION OF CAUSAL DIAMONDS IN (2+1)-DIMENSIONAL GRAVITY

by

Rodrigo Andrade e Silva

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Ted Jacobson, Chair/Advisor
Professor Maissam Barkeshli
Professor Steven Carlip
Professor William Goldman
Professor Raman Sundrum

Preface

This thesis is divided into two (plus one) parts.

In Part I we describe the classical system, the structure of constraints and gauge transformations, and carry out the reduction process to find the reduced phase space of the theory.

In Part II we describe Isham's group-theoretic method of quantization, find an appropriate group to carry the quantization, and discuss some general aspects of the resulting quantum theory.

These two parts, which form the *proper thesis*, are almost verbatim copies of my authorial solo papers, [1] and [2], respectively.

We also add a Part III discussing some perspectives and outlooks, with a much more philosophical and speculative flavor, and consequently less polished and rigorous than the proper thesis.

For a compilation of the main symbols, definitions and conventions used in the text, see Appendix A.

Dedication

To my father José Claudio, my mother Rosmeire and my sister Juliana, who have motivated and supported me always.

Acknowledgments

I am especially grateful to my advisor, Ted Jacobson, for all the enlightening discussions we had, for the immense knowledge and wisdom he shared, for all the support and care he provided, and for his significant contribution in shaping me as a researcher. I also thank him for motivating this project, collaborating on parts of it, and offering valuable feedback.

I thank Luis Apolo, Abhay Ashtekar, Steve Carlip, Laurent Freidel, Marc Henneaux, Jim Isenberg, Alex Maloney, Blagoje Oblak, Pranav Pulakkat, Gabor Sarosi, Antony Speranza, Raman Sundrum and Edward Witten for helpful discussions most directly relevant to this work.

I also thank Stefano Antonini, Batoul Banihashemi, Gong Cheng, Eivind Jørstad, Lucas Nogueira, Pranav Pulakkat, Matheus Schossler, Jacob Taylor, Duy Vu and Yixu Wang for the many interesting interactions.

This research was supported in part by the National Science Foundation under Grants PHY-1708139, PHY-2012139 and PHY-2309634 at the University of Maryland; and in part by the National Science Foundation under Grants No. NSF PHY-1748958 and PHY-2309135 at the Kavli Institute for Theoretical Physics. I am grateful for the hospitality of Perimeter Institute for Theoretical Physics where part of this work was carried out.

Contents

1 Introduction

One of the grand goals of modern physics is to develop a consistent theory of quantum gravity, a problem that has confounded physicists for almost a century and is still unresolved despite intense research. Without the help of experimental evidence to steer us in the right direction, possibly the best course of action is to commit to a set of principles and proceed through logic and mathematical rigor to explore the ultimate consequences of these choices. Here we adopt the principle that quantum gravity is a quantum mechanical theory of gravity, and moreover that it is obtained from a canonical quantization of general relativity. The first principle derives from the perspective that quantum mechanics is based on a very rigid structure (namely, the complex linear structure of operators on Hilbert spaces), which has been tested in many different scenarios with no hint of violation — before adventuring into the exploration of more radical fundamental theories, it is fair to take the conservative stance and simply trust quantum mechanics until further conceptual revision is called for. The second principle is that of canonical quantization, a prescription proposed by Dirac to infer the quantum theory underlying a given classical theory, which has been remarkably successful in many situations. Although it is possible (and even likely) that the "fundamental" theory of quantum gravity does not correspond to a quantization of general relativity, it is still plausible that such a quantization could provide a partial, approximated picture for quantum gravity and yield sufficient insight to motivate the next leap forward in this endeavor.

In this work we develop a non-perturbative canonical quantization of causal diamonds in (2+1)-dimensional gravity. By causal diamonds we mean a class of finite-sized, globally-hyperbolic spacetimes whose Cauchy slices have the topology of a ball — each "diamond state" is defined as the maximal development of Einstein's equation from initial data given on a Cauchy slice. The motivations, which will be further explained in this introduction, are two-fold. The first is that we want to better understand quantum gravity in a quasi-local sense. That is, what is the proper notion of "spacetime subregion" (or even "spacetime" itself) in quantum gravity, given that there are no compactly-supported gauge-invariant observables in gravity and therefore no clear notion of locality? At least from the classical perspective, causal diamonds are the natural object to study as they best represent a finite self-contained subsystem of spacetime. The second motivation is to explore a particular program for canonically quantizing gravity which seems promising in many situations, in arbitrary dimensions, based on the symplectic reduction of the phase space using a convenient gauge-fixing of time by constant-mean-curvature surfaces. If one can prove that this gauge-fixing is well-posed for the class of spacetimes under consideration, the constraints simplify and, and in many cases, can be solved (in principle) yielding a universal characterization of the reduced phase space as the cotangent bundle of the space of conformal geometries on the Cauchy slice. However, for this method to apply to the problem, we need to assume some energy condition (at the classical level) and also some boundary condition. Here we assume pure gravity with a non-positive cosmological constant and "Dirichlet condition" for the metric induced on the boundary of the Cauchy slice. Moreover, to have a better handle on the problem, and particularly on its quantization, we assume 2+1 spacetime dimensions.

This work is divided into two (plus one) parts. In Part I we describe the classical system, the structure of constraints and gauge transformations, and carry out the reduction process to find the reduced phase space of the theory; in Part II we describe Isham's group-theoretic method of quantization, find an appropriate group to carry the quantization, and discuss some general aspects of the resulting quantum theory. These two parts are almost verbatim copies of my authorial solo papers, [1] and [2], respectively. Some of the results are also summarized, in a brief but fairly explicit manner, in [3].[1] In addition to these two parts, which form the *proper thesis*, we also include a Part III with some perspectives and outlooks, displaying a much more philosophical and speculative flavor, and consequently less polished and rigorous.

1.1 Motivations

It has long been known that perturbative canonical quantization of general relativity, expanded around a fixed background geometry, does not lead to a complete theory. In particular, it is not renormalizable since the coupling parameter, G,

[1] While a few aspects of this research were carried out in collaboration with my advisor, and paper [3] is coauthored with him, nearly all of the work represented in the thesis is my own.

has negative mass dimension in spacetime dimensions equal or greater than three.[2] There remains the hope that a careful, non-perturbative quantization of general relativity could still be meaningful. Among the other challenges encountered in the quantization of gravity, one is the famous *problem of time* [5, 6, 7]. As time is a dynamical aspect of gravity, as opposed to a background entity, its role and place in the quantum theory is enigmatic. Another issue is the non-linearity of the constraints of general relativity, particularly in the momentum variables, which severely complicates attempts to proceed with the quantization exactly and non-perturbatively. Lastly, a notable peculiarity of gravity is the absence of local, or even compactly supported, observables. This is because any physical observables must be invariant under gauge transformations, and therefore must be invariant under general spacetime diffeormorphisms (at least those that are supported away from the boundary). Thus, the general notion of locality, and even the meaning of subregions, is particularly fuzzy in quantum gravity.

In view of these general challenges, we attempt to the address the following two main points in this work:

1. We wish to better understand how to describe quantum gravity in finite regions of space(time). In particular, what can be learned if we take the classi-

[2] Notwithstanding this power-counting argument, gravity in three spacetime dimensions is actually perturbatively renormalizable, which can be seen when expressed as a topological Chern-Simons theory [4]. For a closed Cauchy slice this is expected since the theory has a finite number of degrees of freedom, but in asymptotically AdS or in the presence of spatial boundaries the conclusion may not be as clear—in fact, since the theory has no local degrees of freedom, the question of pertubative renormalizability, in the context of local quantum field theories, might be ill-posed.

cal notion of a self-contained subregion of spacetime, i.e. a *causal diamond*, and quantize (Einstein-Hilbert) gravity inside it. More precisely, we wish to quantize the class of spacetimes consisting of causal diamonds in pure general relativity (including a cosmological constant), where a causal diamond is defined as the maximal development of initial data, satisfying the constraints, given on a bounded acausal spatial slice.

2. We wish to continue the exploration of a program for quantizing gravity non-perturbatively by explicitly reducing the phase space via a particular gauge-fixing for time defined by a CMC (constant-mean-curvature) condition [8, 9, 10]. If one is quantizing a class of spacetimes in which each spacetime admits a regular CMC foliation (i.e., where the leaves are defined by having a constant trace of extrinsic curvature, a.k.a, mean-curvature), then the constraints of general relativity can be cast into a more manageable form, in terms of the *Lichnerowicz equation*; if one can prove certain existence and uniqueness properties for the solution of this equation, then quite generally the reduced phase space (of pure gravity) ends up being $T^*[\mathrm{ConGeo}(\Sigma)]$, i.e., the cotangent bundle of the space of conformal geometries[3] on the Cauchy slice. This yields a non-perturbative characterization of the reduced phase space, which completely resolves the issues related to the constraints and gauge in-

[3]The space of conformal geometries, $\mathrm{ConGeo}(\Sigma)$, is the space of equivalence classes of metrics h on a manifold Σ where two metrics are identified if they can be related by a combined Weyl scaling and diffeomorphism push-forward, $h \sim \Psi_*\Omega h$; the class of metrics h, positive functions Ω and diffeomorphisms Ψ participating in this quotient depends on the details of gravitational system being reduced.

variance of gravity at the classical level; accordingly, it is in principle easier to take the next step and try to quantize the resulting (reduced) theory non-perturbatively.

Another reason for attempting to combine these two points is that certain spacetimes do not admit CMC foliations (or even a maximal slice), due to global reasons. For example, any non-flat spacetime with topology $T^3 \times \mathbb{R}$, satisfying the timelike convergence condition (i.e., $\boldsymbol{Ric}(u, u) \geq 0$ for all timelike vectors u), does not posses a maximal slice [11, 12, 13, 14, 15], and therefore cannot be treated with this Lichnerowicz method. Even if there is a foliation by CMCs, the Lichnerowicz equation may not have the desired existence and uniqueness properties (e.g., see [16]). If one could properly quantize local regions of spacetime via this Moncrief-Lichnerowicz approach, and glue the resulting "quantum causal diamonds" to form a larger spacetime, then one would have a theory of quantum gravity describing a much larger class of spacetimes.

While promising in principle, the actual implementation of this program for general causal diamonds, in arbitrary dimensions, would be a quite ambitious effort. First, it would involve a classification of the space of conformal geometries in n-dimensional manifolds with boundaries, which is not a fully solved mathematical problem; second, one would have to develop a non-perturbative quantization of such space, which would presumably be challenging. In view of the difficulties, it is worthwhile to study the problem in a simplified setting, where we might have a better handle on both the physics and the mathematics. In this work, the problem of interest is 2+1 dimensional (Einstein-Hilbert) gravity with a nonpositive cosmological constant in the domain of dependence of a topological disc, hereinafter simply referred to as a *causal diamond*. For reasons that will be explained later, the class of spatial metrics are restricted by a "Dirichlet boundary condition", that is, the (induced) metric is fixed at the boundary of the disc. We reiterate that this is not a theory of a single causal diamond, but rather a dynamical theory of gravity in the class of globally-hyperbolic spacetimes whose Cauchy slices are topological discs (with fixed induced corner metric). The reason for considering a nonpositive cosmological constant is that it ensures that the spacetime can be entirely foliated by CMC surfaces, thus providing a natural notion of time for causal diamonds, and allowing the application of the CMC gauge-fixing of time, and also that the associated Lichnerowicz equation can be proven to have the desired existence and uniqueness properties.

There exists an extensive literature on (2+1)-dimensional gravity systems. A limited sample of references includes work on spacetimes with closed spatial slices (where the reduced phase space is finite-dimensional) [4, 17, 9, 8, 10, 18, 19, 20, 21, 22], on spacetimes with finite timelike boundary [23, 24, 25], and on asymptotically AdS_3 spacetimes[26, 27, 28, 29, 30, 31, 32, 33]. The study of causal diamonds can be valuable to improve the understanding of subsystems or "regions of spacetime' in quantum gravity — as mentioned above, the dynamical nature of spacetime and the diffeomorphism invariance of the theory, which precludes the existence of local (or quasi-local) observables, makes the notion of "subregions of spacetime" particularly fuzzy — deservedly, they have received a great deal of attention recently from a variety of different approaches attempting to unveil their quantum properties [34, 35, 36, 37, 38]. In this work we intend to push the analysis further by developing a fully non-perturbative quantization of (pure) gravity in causal diamonds spacetimes, via the phase space reduction approach paired with Isham's group-theoretic quantization. While simple enough to be exactly solvable classically, due to the absence of local degrees of freedom, the system has nevertheless an infinite-dimensional reduced phase space of "boundary gravitons". (A recent paper by Witten [39] revitalizes this program for canonically quantizing gravity, applying a similar approach to asymptotically Anti-de Sitter spacetimes.)

1.2 Summary

We begin with a somewhat detailed summary of the contents of the thesis. The goal is to provide a quick guide to the main ideas and results, in view of the extensive nature of the paper.

Note: Appendix A contains a compilation of the main symbols, definitions and conventions used in the text.

1.2.1 Summary of Part I

In quantizing a gauge theory one must, sooner or later, deal with the constraints and gauge invariance of the theory. There are two mainstream routes of proceeding. The first route, called the Dirac approach, is to first quantize the theory ignoring the constraints, and then impose the constraints at the quantum level. That is, if the (unconstrained) phase space is covered by conjugate coordinates q^i and p_i, and there is a set of constraints $C_\alpha = 0$, the quantization goes by first constructing a Hilbert space carrying a representation of the Heisenberg algebra $[\widehat{q}^i, \widehat{p}_j] = i\hbar\delta^i_{\ j}$, and then restricting to the *physical* subspace defined by states satisfying the conditions $\widehat{C}_\alpha|\psi\rangle = 0$, and to physical, gauge-invariant observables that commute with the constraints, $[\mathcal{O}, \widehat{C}_\alpha] = 0$.[4] The second route, called the reduced phase space approach (Sec. 2), is to first impose the constraints at the classical level, remove all the associated gauge ambiguities, and then quantize the resulting (gauge-free) theory. There is no a priori guarantee that the two approaches to quantization would lead to the same quantum theory. Here, for the quantization of causal diamonds, we shall focus on the reduced phase space approach.

The gravitational system of our interest is defined (Sec. 3.1) as the class of maximal developments, under the vacuum Einstein's equations (with a non-positive cosmological constant), of all initial geometric data (satisfying a certain boundary condition) given on a manifold with the topology of a 2-dimensional disc (with

[4] It may be necessary to only impose these constraints weekly, i.e., in between physical states.

boundary). According to the ADM (Arnowitt-Deser-Misner) formalism, the geometrical data consists of a spatial metric h_{ab} and an extrinsic curvature K^{ab}, which must satisfy the momentum constraint $\nabla_a(K^{ab} - Kh^{ab}) = 0$ and the Hamiltonian constraint $K^{ab}K_{ab} - K^2 - R_{(h)} + 2\Lambda = 0$, where $K := K^{ab}h_{ab}$ is the trace of the extrinsic curvature and $R_{(h)}$ is the Ricci scalar for h_{ab}. Each given geometric data defines a *causal diamond*. We reiterate that we will be quantizing not a single causal diamond, but the class of all such causal diamonds satisfying a certain condition on the corner (i.e., the boundary of any Cauchy slice) metric. Note that a typical causal diamond in this class is not the intersection of the past and future of two timelike-separated points, but rather it will have horizons with a ridged, mountain-like appearance, as in Fig. 1. In general the momentum constraint is the simpler

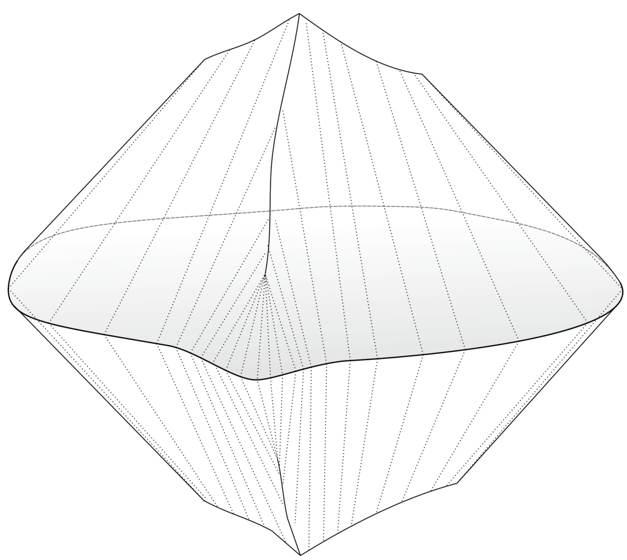


Figure 1: A typical causal diamond, obtained by maximally developing geometric data on a Cauchy slice with disc topology.

one to solve, since it is linear in the momentum variable; the trouble in general relativity is solving the Hamiltonian constraint. We implement a program (Sec. 3.2) based on the gauge-fixing of time by surfaces of constant mean-curvature. That is, given any foliation of a causal diamond, one uses the gauge flow generated by the Hamiltonian constraint to re-foliate the diamond into slices of constant K, which varies monotonically along the foliation and thus define a suitable time coordinate $\tau = -K$. As K is now a time-dependent constant, on each spatial slice, only the trace-free part of K^{ab}, $\sigma^{ab} := K^{ab} - \frac{1}{2}Kh^{ab}$, remains as a dynamical variable together with h_{ab}. In this gauge, if one starts with some "seed data", (h_{ab}, σ^{ab}), which is taken to satisfy (at least) the momentum constraint, which now reads $\nabla_a \sigma^{ab} = 0$, then it may be possible to deform this seed data into "valid data", i.e. so that both the momentum and Hamiltonian constraints are satisfied, via an appropriate Weyl transformation, $(h_{ab}, \sigma^{ab}) \mapsto (e^{\phi}h_{ab}, e^{-2\phi}\sigma^{ab})$; the condition for this deformed data to satisfy the constraints is that ϕ must satisfy a non-linear elliptic equation called the *Lichnerowicz equation*.

Three things must be checked in order to establish the non-perturbative validity of this (partial) gauge-fixing prescription, and a fourth one to ensure utility. First, one needs to ensure that the CMC foliation always exists and is unique for every causal diamond within the class of causal diamonds under consideration. We will argue (Sec. 4.1) that the foliation exists and covers each diamond entirely as τ ranges from $-\infty$ to ∞, provided that $\Lambda < 0$; the case $\Lambda = 0$ can be included by a continuity argument. The second point is that the CMC foliation can be attained by a gauge transformation, starting from any other foliation (Sec. 4.2). It is thus important to make sure that generic smearings of the Hamiltonian constraint, $H[\eta] := \int_\Sigma d^2x\sqrt{h}\,\eta(K^{ab}K_{ab} - K^2 - R_{(h)} + 2\Lambda)$, where η vanishes at the boundary $\partial\Sigma$ (since all Cauchy slices Σ in a causal diamond meet at the corner), must be gauge-generators. As the constraints are first-class, $H[\eta]$ is a generator of gauge provided that it produces a well-defined flow in the (pre) phase space. With respect to the symplectic form $\Omega = \int_\Sigma d^2x\,\delta\pi^{ab} \wedge \delta h_{ab}$, where $\pi^{ab} = \sqrt{h}(K^{ab} - Kh^{ab})$, the flow of a phase space function is regular if and only if it has well-defined functional derivatives, i.e., if $\delta H[\eta]$ is of the form $\int_\Sigma d^2x(A^{ab}\delta h_{ab} + B_{ab}\delta\pi^{ab})$. In the present case, one notices that for arbitrary boundary conditions, only the $H[\eta]$ with vanishing normal derivative of η at the boundary are gauge generators; in other words, corner boosts (which tilt the angle at which the Cauchy slice meets the corner) are not gauge transformations, but rather non-trivial transformations between different states. However, if one imposes a "Dirichlet boundary condition", where the induced metric at the boundary is fixed, then all $H[\eta]$, with $\eta|_\partial = 0$, are gauge-generators. Therefore we restrict the class of spatial metrics in this way, so that the CMC gauge is attainable. Third, one needs to ensure that the Hamiltonian constraint can be solved by a Weyl transformation, that is, for any seed data (h_{ab}, σ^{ab}) there must exist a solution ϕ of the associated Lichnerowicz equation. We show that the equation always has solutions, for all τ, as long as $\Lambda \leq 0$ (Sec. 4.3). Fourth, in determining the reduced phase space, it is important that there are no residual, unfixed gauge directions; in this language, one needs to prove that there is a unique solution ϕ for any given seed data. This can be shown to also follow from $\Lambda \leq 0$, together with the fact that the Dirichlet boundary condition requires $\phi|_\partial = 0$, as we consider only seed data that satisfies the boundary condition (Sec. 4.3). Uniqueness is essential since it implies that when two seed data related by a Weyl transformation, $(h', \sigma') = (e^\lambda h, e^{-2\lambda}\sigma)$, are used as inputs in the Lichnerowicz algorithm, they will each be deformed into the same output valid data. This means that the constraint surface (i.e., the set of all valid data) can be identified with the set of equivalence classes of seed data under Weyl transformations (acting trivially at the boundary).

Up to this point we have dealt with the gauge associated with time refoliations, and solved the constraints; we need next to deal with gauge associated with spatial diffeomorphisms. As one could anticipate, it is possible to show that two valid data that can be related by a diffeomorphism that is trivial at the boundary, $(h'_{ab}, \sigma'^{ab}) = (\Psi_* h_{ab}, \Psi_* \sigma^{ab})$, where $\Psi : \Sigma \to \Sigma$ satisfies $\Psi|_\partial = I$, are gauge related. The reduced phase space is the space of physically distinguishable solutions to the equations of motion, or equivalently the space of valid initial data modulo gauge transformations. Thus, here it can be identified with the set of equivalence classes of valid data under these "boundary-trivial" spatial diffeomorphisms, or equivalently the set of equivalence classes of seed data under boundary-trivial conformal transformations. The reduction process can be summarized in a diagram, Fig. 2.

The reduced phase space, just identified with the space of seed data modulo

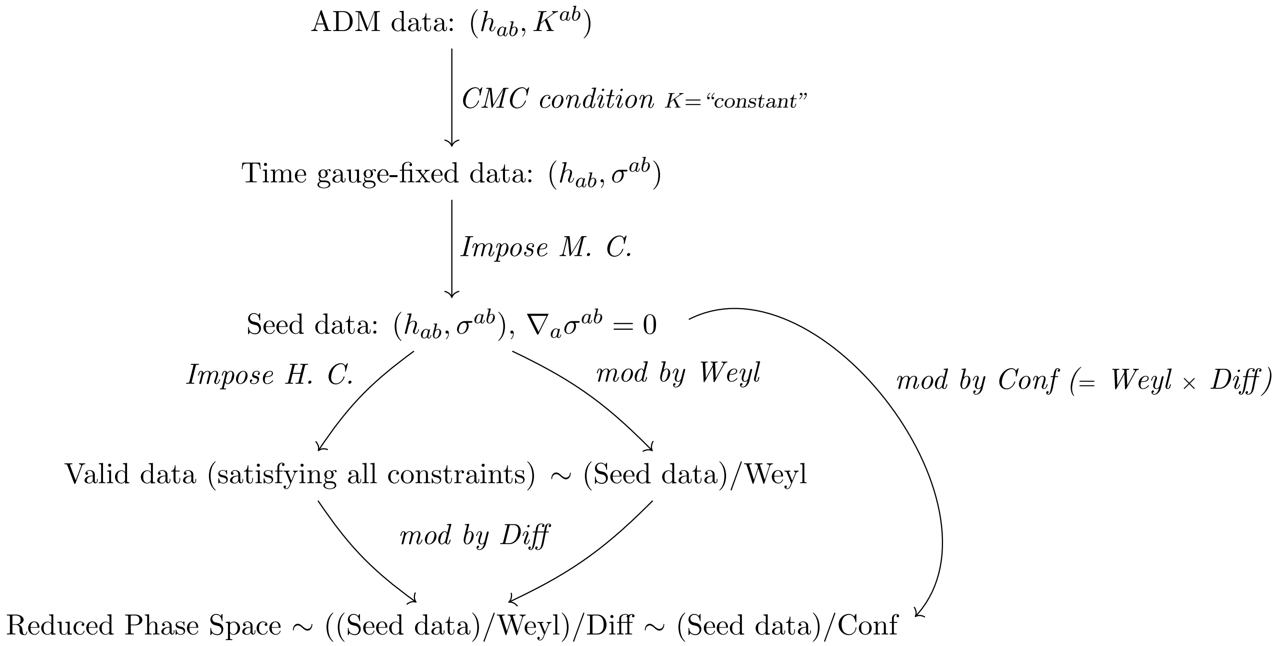


Figure 2: Starting with ADM data, satisfying the Dirichlet condition on the (induced) boundary metric, we gauge-fix "time" by imposing the CMC condition. Then we impose the momentum constraint (M. C.) to define "seed data". Further imposing the Hamiltonian constraint (H. C.) leads to "valid data" (i.e., data satisfying all the constraints); from arguments involving the Lichnerowicz equation, the space of valid data can be identified with the space of seed data modulo (boundary-trivial) Weyl transformations. The reduced phase space is then obtained by further quotienting the space of valid data by (boundary-trivial) diffeomorphisms; this identifies the reduced phase space with the space of seed data modulo (boundary-trivial) conformal transformations.

conformal transformations, can also be identified with the *cotangent bundle of conformal geometries* on the Cauchy slice (Sec. 5.1), with its natural symplectic form. To get a "taste of why" note that, locally in the space of metrics, the infinitesimal directions that will be quotiented out by conformal transformations are $h_{ab} \mapsto (1+\lambda)h_{ab} = h_{ab} + \lambda h_{ab}$ and $h_{ab} \mapsto h_{ab} + \pounds_\xi h_{ab} = h_{ab} + 2\nabla_{(a}\xi_{b)}$; then note that σ^{ab} in the seed data is characterized by two properties, tracelessness and divergencelessness, and consequently σ^{ab} can naturally be thought of as 1-form on the cotangent space of ConGeo(Σ) at $[h] = [\Psi_* \Omega h]$ since the pairing $\langle \sigma, \delta h \rangle := \int_\Sigma d^2x \sqrt{h}\, \sigma^{ab} \delta h_{ab}$ is insensitive to the directions that will be projected out, namely $\langle \sigma, \lambda h_{ab} \rangle = 0$ and $\langle \sigma, \pounds_\xi h_{ab} \rangle = 0$.

Particularizing to the case where Σ is a disc, D, with the Dirichlet condition on the (induced) boundary metric, we can show (Sec. 5.2) that the space of conformal geometries is $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, i.e., the group of (orientation-preserving) diffeomorphisms of the circle modded by the three-dimensional $PSL(2,\mathbb{R})$ subgroup of projective special linear transformations in two real dimensions. (At the algebra level, $\mathfrak{diff}(S^1)$ can be identified with the space of vector fields on S^1, and the $\mathfrak{psl}(2,\mathbb{R})$ subalgebra then corresponds to the three lowest "Fourier modes", ∂_θ, $\sin\theta\,\partial_\theta$ and $\cos\theta\,\partial_\theta$.) The reduced phase space is therefore

$$\widetilde{\mathcal{P}} = T^*\big(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})\big) \tag{1.1}$$

with the natural symplectic form associated with the cotangent bundle structure.

We also present another method for carrying out the reduction of the phase space based on a suitable "choice of coordinates" on the (pre) phase space (Sec. 6). Inspired by the previous argument, we note that by starting with "conformal coordinates" one can immediately detect the null directions of the symplectic form, and then identify the reduced phase space as a "level-set" in these coordinates. In this part we are more rigorous in discussing gauge transformations in terms of the degenerate directions of the symplectic form. This analysis is particularly relevant when there are boundaries, since transformations that would naively be gauge in the bulk can become physical symmetries (mapping between distinct physical states) when acting non-trivially in a neighborhood of the boundary — in particular, we justify that only diffeomorphisms that act trivially on the boundary are gauge. An advantage of this alternative approach is that it provides a natural "coordinatization" of the reduced phase space, along with the explicit quotient map from the redundant but concrete geometrical variables (namely, metrics and extrinsic curvatures) to the physical but abstract variables describing the reduced phase space. This is useful because what characterizes a quantum theory is not the structure of the Hilbert space itself, but the way in which meaningful physical observables are represented in the Hilbert space.[5] When we come to the quantization, the variables in the reduced phase space will be those to become (self-adjoint) operators in the Hilbert space; it is therefore essential that we can understand their physical meaning, i.e., what experimental measurement would be described by that particular operator. Having this explicit map from the abstract variables back to concrete, geometrical variables is therefore valuable in this program.

The conformal coordinates are defined using the fact that, on a disc, any two metrics are conformally-equivalent (Sec. 6.1). That is, given any reference metric $\overline{h}_{ab}$,

[5] Recall that any two separable Hilbert spaces, as typically considered in physics, are isomorphic. One can take a countable orthonormal basis $\{\Psi_n\}$ in one Hilbert space and another countable orthonormal basis $\{\Phi_n\}$ in the other Hilbert space, and simply define a unitary map between them by $\Psi_n \mapsto \Phi_n$. Therefore the Hilbert space of a simple Harmonic oscillator is the same as the Hilbert space of the Standard Model of particle physics.

any other metric h_{ab} can be obtained via a conformal transformation, $h_{ab} = \Psi_* \Omega \overline{h}_{ab}$, for some diffeomorphism Ψ and Weyl factor Ω (which are generally not boundary-trivial). We can then "pull-back" the traceless and divergenceless (with respect to h_{ab}) extrinsic curvature σ^{ab} through this conformal map as $\overline{\sigma}^{ab} := \Omega^2 \Psi^* \sigma^{ab}$. In this way, $\overline{\sigma}^{ab}$ is traceless and divergenceless with respect to the reference metric $\overline{h}_{ab}$. Instead of using geometric variables (h_{ab}, σ^{ab}) to describe the seed data, we can now use *conformal coordinates* $(\Psi, \Omega, \overline{\sigma}^{ab})$. The Dirichlet condition on the induced boundary metric fixes the boundary value of Ω in terms of the boundary action of Ψ, $\psi := \Psi|_\partial$, and the Lichnerowicz equation then fixes uniquely the entire value of Ω in terms of ψ and $\overline{\sigma}^{ab}$; thus the constraint surface is parametrized (Sec. 6.2) only by $(\Psi, \overline{\sigma}^{ab})$. At this stage, by looking at the symplectic form, it becomes evident that changing $\Psi \mapsto \Psi'$ is a gauge transformation as long as $\Psi|_\partial = \Psi'|_\partial$; which leads (Sec. 6.3) to a partial phase space reduction $(\Psi, \overline{\sigma}^{ab}) \mapsto (\psi, \overline{\sigma}^{ab})$. The fact that $\overline{\sigma}^{ab}$ is traceless and transverse (with respect to $\overline{h}_{ab}$) implies that it can be described by boundary data. More precisely, the space of such $\overline{\sigma}^{ab}$ is naturally isomorphic to the subspace of $\mathfrak{diff}^*(S^1)$ (the dual Lie algebra of *Diff*$^+(S^1)$) that annihilates the $\mathfrak{psl}(2, \mathbb{R})$ subalgebra of $\mathfrak{diff}(S^1)$; this space will be denoted by $\mathring{\mathfrak{diff}}^*(S^1)$, and its elements by $\mathring{\sigma}$. At this stage, the phase space is thus described by $(\psi, \mathring{\sigma}) \in \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$. But by direct inspection of the symplectic form we find that there are still three null directions per phase space point. The component of these directions on the *Diff*$^+(S^1)$ factor are directly related to $PSL(2, \mathbb{R})$, but they also have a non-trivial component along the $\mathring{\mathfrak{diff}}^*(S^1)$ factor (Sec. 6.4). We construct explicitly a map J that quotients out those directions, leading to the fully reduced phase space $\widetilde{\mathcal{P}} = T^*\big(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})\big)$.

Lastly we study the Hamiltonian generating time-evolution between CMC slices (Sec. 7). We first review some basic facts about determining the Hamiltonian in a gauge-fixed, reduced phase space approach (Sec. 7.1). In a pure reduction process, the Hamiltonian on the reduced phase space is simply obtained by "pushing-forward" the original Hamiltonian. This push-forward is well-defined because, in a consistently formulated dynamical theory, the Hamiltonian is gauge-invariant and therefore constant within the pre-image (under the quotient map) of any point of the reduced phase space. When there is a gauge-fixing involved, the original and reduced Hamiltonians generating evolution along the "time parameter" are not related so simply. The easiest manner to determine the reduced Hamiltonian is by looking at the action (Sec. 7.2), from which we recover the expected result [40] that the Hamiltonian generating evolution between CMC slices is given, at time τ, by the area of the slice with $K = -\tau$. Despite its simple geometrical interpretation, its expression in terms of reduced phase space variables is highly non-trivial since it is defined implicitly with respect to solutions of the Lichnerowicz equation. In fact, it is only evident from its formula that it is a well-defined function on the partially-reduced phase space, $\mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$. As a check of consistency (Sec. 7.3), we show explicitly that this Hamiltonian is indeed a well-defined function on

$T^*\big(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})\big)$. Because this Hamiltonian is a complicated function on the reduced phase space, it is interesting to explore some regimes where it can be solved analytically, or at least approximated (Sec. 8). This is relevant if one wishes to describe the quantum dynamics of the system. We stress that while the full classical reduction and the kinematical part of the quantization (i.e., representing a complete algebra of observables on a Hilbert space) are performed non-perturbatively, the dynamics (in CMC time) may well only be only amenable to approximate analysis.

Other than App. A — *glossary, symbols and conventions* — there are two more appendices. Appendix B is a review of the Riemann mapping theorem, a special case of the uniformization theorem, which shows that all (Riemannian) metrics on a disc are conformally-equivalent. Appendix C describes the embedding picture, explaining how to construct a map from a point in the reduced phase space to a causal diamond embedded in AdS_3 (or $Mink_3$ if $\Lambda = 0$). For intuition and artistic purposes, some accurate pictures of causal diamonds are displayed.

1.2.2 Summary of Part II

By solving all the constraints of general relativity and eliminating the associated gauge ambiguities, we have shown in Part I that the the fully reduced phase space for the causal diamonds, with fixed boundary metric, is $\widetilde{\mathcal{P}} = T^*\big(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})\big)$, with the natural symplectic form coming from the cotangent bundle structure. The problem then becomes quantizing a non-gauge theory. Canonical quantization is a remarkable tool discovered by Dirac that is based on the principle that the classical theory and its corresponding quantum theory share (roughly) the same underlying algebraic structure of observables (Sec. 9.1). In its original and simplest incarnation, one takes a complete set of canonically conjugate coordinates x^i and p_j, satisfying the Poisson algebra $\{x^i, p_j\} = \delta^i{}_j$, and then defines the quantum theory in such a way that self-adjoint operators X^i and P_j form an irreducible representation, on a Hilbert space, of the analogous algebra $\frac{1}{i\hbar}[X^i, P_j] = \delta^i{}_j$, where $\hbar$ is the Planck constant (which has dimensions of angular momentum, like $[X, P]$). This rule has worked magnificently in many different scenarios, from the quantization of Newtonian particles in Euclidean space to relativistic fields in Minkowski space. There is, however, an important limitation: it only produces a sensible quantum theory if the phase space has a natural linear structure, $\mathcal{P} = \mathbb{R}^{2n}$, so that there is a natural class of global coordinates x^i and p_j, ranging from $-\infty$ to $+\infty$, which can be used to produce a preferred quantum theory. In particular, notice that from the assumption that P is self-adjoint it follows that $e^{aP/i\hbar}$ is a well-defined, bounded, unitary operator for any $a \in \mathbb{R}$; and if X has an eigenvector $|\psi\rangle$ with eigenvalue $x \in \mathbb{R}$, then it follows from the algebra $[X, P] = i\hbar$ that $e^{aP/i\hbar}|\psi\rangle$ is another eigenvector of X with eigenvalue $x + a$; therefore, if the spectrum of X (and similarly that of P) is not empty (i.e., the theory is non-trivial), then it must be the whole real line. For this reason, when the phase space has a non-trivial topology, or lacks a natural global chart of "Cartesian coordinates", a more sophisticated method of quantization is necessary.

In our case, while $T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$ happens to be topologically contractible, it appears not to have a natural linear structure. Opportunely, it does admit a natural group action, as the configuration space $\mathcal{Q} := \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ is a homogeneous space for the group $\mathit{Diff}^+(S^1)$, so it is reasonable to employ Isham's group-theoretic method of quantization [41, 42]. In this method (Sec. 9.2), one takes a transitive[6] group $\widetilde{G}$ of symplectic symmetries of the phase space and uses it to define a set of classical observables and their corresponding quantum operators, where the Hilbert space is constructed to carry a (projective) irreducible unitary representation of the group. The justification comes from the core principles of canonical quantization: any element ξ in the Lie algebra $\widetilde{\mathfrak{g}}$ of $\widetilde{G}$ generates a Hamiltonian flow X_ξ on the phase space, to which is associated a Hamiltonian charge H_ξ (solution of $dH_\xi = -i_{X_\xi}\omega$); given any basis ξ_i of $\widetilde{\mathfrak{g}}$, $i = 1 \ldots \dim(\widetilde{\mathfrak{g}})$, we have a set of classical charges $H_i := H_{\xi_i}$ whose Poisson algebra is homomorphic to $\widetilde{\mathfrak{g}}$, i.e., $[\xi_i, \xi_j] = c^k_{ij}\xi_k \Rightarrow \{H_i, H_j\} = c^k_{ij}H_k$ (we assume that the group has been extended, if necessary, to incorporate a non-trivial central charge that could otherwise subvert this homomorphism); moreover, the transitivity of $\widetilde{G}$ implies that this set of charges is complete in the sense that the specification of their value determines the phase space point (up to, possibly, a discrete ambiguity); therefore, the underlying group of symmetries ensures that there is a well-grounded, complete and algebraically-closed set of classical charges on the phase space, which can then be quantized according to

[6] A group G acts transitively on a manifold $\mathcal{M}$ if for any two points x, $x' \in \mathcal{M}$ there exists $g \in G$ such that $x' = gx$.

the traditional rule, i.e., $H_i \mapsto \widehat{H}_i$ where $\widehat{H}_i$ are self-adjoint operators on a Hilbert space forming an irreducible representation of the algebra $\frac{1}{i\hbar}[\widehat{H}_i, \widehat{H}_j] = c_{ij}^k \widehat{H}_k$. In this context, $\widetilde{G}$ is called the *canonical group*, and the H's are *canonical charges*.

As an example, in the case of a particle on the Euclidean line (i.e., phase space $\mathcal{P} = T^*\mathbb{R}$), notice that p generates translations in x while x generates translations in $-p$. So instead of thinking of the coordinates x and p as the elementary ingredients for quantization, reverse the logic and think of this $\mathbb{R}^2$ group of symmetries, $(x, p) \mapsto (x + a, p - b)$, as the starting point, i.e., where now x and p are *derived* from the group as the associated canonical charges. In this case, the corresponding Poisson algebra does introduce a non-trivial central charge, $\{x, p\} = 1$, implying that we need to consider a central extension of $\mathbb{R}^2$, namely the Heisenberg group $H(3) = \widehat{\mathbb{R}^2}$. Accordingly, the quantum theory is constructed in terms of a unitary irreducible representation of $\widetilde{G} := H(3)$, which is unique (up to unitary equivalence) and precisely the familiar one realized by square-integrable $\mathbb{C}$-valued wavefunctions on $\mathbb{R}$.

To construct a natural canonical group for the causal diamonds we start with the group $G = \mathit{Diff}^+(S^1)$ acting on the configuration space $\mathcal{Q} = \mathit{Diff}^+(S^1)/\mathit{PSL}(2, \mathbb{R})$ from the left as $\delta_\phi[\psi] := [\phi \circ \psi]$, where ϕ, $\psi \in \mathit{Diff}^+(S^1)$ and $[\psi] \in \mathcal{Q}$. This action is transitive on $\mathcal{Q}$ and naturally lifts to a G-action $\widetilde{\delta}_\phi$ on the cotangent bundle $\widetilde{\mathcal{P}} = T^*\mathcal{Q}$, in which 1-forms p at $[\psi]$ are pulled-back by ϕ^{-1} to a 1-form at $[\phi\psi]$, i.e., $\widetilde{\delta}_\phi p := \delta^*_{\phi^{-1}} p$. While this action indeed defines a symplectic symmetry of $\widetilde{\mathcal{P}}$, it is not transitive on $\widetilde{\mathcal{P}}$ because it does not move points "vertically" (i.e., along the fibers of $T^*\mathcal{Q}$). There is a general prescription (Sec. 9.3) to extend such a group G, acting transitively on $\mathcal{Q}$, to a group $\widetilde{G}$ acting transitively on $T^*\mathcal{Q}$ as symplectomorphisms. The core step is to find a representation of G on some vector space V with the property that at least one of the G-orbits is diffeomorphic to $\mathcal{Q}$. In this setting, every dual vector $\alpha \in V^*$ defines a "momentum translation" as follows: as $\mathcal{Q}$ is embedded in V, tangent vectors to $\mathcal{Q}$ can be identified with vectors v in V, and thus α defines a 1-form field on $\mathcal{Q}$ mapping $v \in T\mathcal{Q}$ to $\alpha(v) \in \mathbb{R}$; this 1-form field can then be used to translate $p \in T^*\mathcal{Q}$ as $p \mapsto p - \alpha$, which thus defines a group action of V^* on $\widetilde{\mathcal{P}}$. This vertical action is a symplectomorphism, and most importantly the combined G and V^* actions form a transitive group $\widetilde{G} := V^* \rtimes G$ of symplectomorphisms of the phase space $T^*\mathcal{Q}$.

With this general machinery at hand, we can return to our causal diamonds (Sec. 10). Despite the apparent promise of starting with the $\mathit{Diff}^+(S^1)$ action, we could not find a linear representation of it with the required property of having an orbit diffeomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. Fortunately, it happens that the coadjoint representation of the Virasoro group does have an orbit diffeomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. The Virasoro group, being a central extension of $\mathit{Diff}^+(S^1)$ by $\mathbb{R}$, $\mathit{Vira} = \widehat{\mathit{Diff}^+(S^1)}$, can just as well be used as the group of "configuration translations", $G = \mathit{Vira}$, where the central element acts trivially on the configuration space, i.e., $\delta_{(\phi,a)}[\psi] := [\phi \circ \psi]$. That is, the configuration space can be equivalently expressed as the homogeneous space $\mathcal{Q} = \mathit{Vira}/(PSL(2,\mathbb{R}) \times \mathbb{R})$. For completeness and compatibility of notation, we review some basic facts about the Virasoro group (Sec. 10.1), and properties of its adjoint (Sec. 10.1.1) and coadjoint (Sec. 10.1.2) representations, with particular interest in describing how $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ is embedded as a coadjoint orbit of Virasoro into $V = \mathfrak{vira}^*$ (the dual Lie algebra of Vira). Thus, we propose (Sec. 10.2) that the quantization of causal diamonds should be based on the canonical group that is the semi-direct product

$$\widetilde{G} = V^* \rtimes G = (\mathfrak{vira}^*)^* \rtimes \mathit{Vira} \tag{1.2}$$

where Vira acts as "configuration translations" (i.e., lifted from the action on the configuration space, thus moving points "laterally" on $T^*\mathcal{Q}$) and the abelian normal subgroup $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$ acts as "momentum translations" (i.e., moving points vertically on $T^*\mathcal{Q}$). In more detail, this group acts on the phase space as follows: when $\mathcal{Q}$ is realized as a coadjoint orbit of Virasoro, a point $[\psi] \in \mathcal{Q}$ is identified with a vector in $\mathfrak{vira}^*$; a point p in phase space (which is, say, a 1-form at $[\psi]$) is identified with a pair $(\widehat{\rho}, [\psi])$, where $\widehat{\rho}$ is some dual vector in $\mathfrak{vira}^*$ (i.e., an element of $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$); then a group element $(\widehat{\xi}; \widehat{\phi}) \in \mathfrak{vira} \rtimes \mathit{Vira}$ acts on p as $\Gamma_{(\widehat{\xi}, \widehat{\phi})}(\widehat{\rho}, [\psi]) = (\mathrm{ad}_{\widehat{\phi}}\widehat{\rho} - \widehat{\xi}, [\phi\psi])$, where ad denotes the adjoint action of Vira on $\mathfrak{vira}$. To close this section, we describe (Sec. 10.3) how to lift the group action to the partially-reduced phase space $\widehat{\mathcal{S}} = \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$ (whose structure is simpler to manipulate) and also to the constrained ADM phase space $\mathcal{P}$ (characterized in terms of geometrical variables, which can therefore be useful in extracting the physical meaning of the symmetries and the associated charges).

The next step (Sec. 11) is to evaluate the canonical charges associated with $\widetilde{G} = \mathfrak{vira} \rtimes Vira$, denoted by $H_{(\widehat{\eta};\widehat{\xi})}$, where $(\widehat{\eta};\widehat{\xi}) \in \widetilde{\mathfrak{g}} = \mathfrak{vira}^c \oplus \mathfrak{vira}$ (in which $\mathfrak{vira}^c$ denotes the algebra of the abelian group $\mathfrak{vira}$, i.e., $\mathfrak{vira}^c$ is isomorphic to $\mathfrak{vira}$ as a vector space but has a commutative algebraic structure). This can be obtained from the general formulas derived in Sec. 9.3 for phase spaces with a cotangent bundle structure. The *Vira* part of $\widetilde{G}$ act as configuration (or lateral) translations and therefore the corresponding charges are interpreted as "momentum variables", denoted by $P_{\widehat{\xi}} := H_{(0;\widehat{\xi})}$; and the (abelian) $\mathfrak{vira}$ part of $\widetilde{G}$ act as momentum (or vertical) translations so the corresponding charges are interpreted as "configuration variables", denoted by $Q_{\widehat{\eta}} := H_{(\widehat{\eta};0)}$. These charges can be decomposed in a convenient Fourier basis: note that elements $\widehat{\xi}$ of $\mathfrak{vira}$ are characterized by a vector field on S^1 plus a central component, $\widehat{\xi} = \xi(\theta)\partial_\theta + x\widehat{c}$, and the vector field can be expanded in a Fourier series, $\xi(\theta) = \sum_{n\in\mathbb{Z}} \xi_n e^{in\theta}$; the canonical charges can then be decomposed in terms of momentum modes $P_n := H_{(0;e^{in\theta}\partial_\theta)}$, a momentum central charge $P_R := H_{(0;\widehat{c})}$, configuration modes $Q_n := H_{(e^{in\theta}\partial_\theta;0)}$ and a configuration central charge $Q_T := H_{(\widehat{c};0)}$. Also, the additive constant ambiguities in the definition of the canonical charges can be chosen so that no additional central charges appear in the Poisson algebra, in which we take $P_R = 0$ and $Q_T = 1$. The Poisson algebra of the canonical charges then becomes

$$\{P_n, P_m\} = i(n-m)P_{n+m}$$
$$\{Q_n, P_m\} = i(n-m)Q_{n+m} - 4\pi i n^3 \delta_{n+m,0}$$
$$\{Q_n, Q_m\} = 0 \tag{1.3}$$

Notably, this corresponds to a $\mathfrak{bms}_3$ algebra, which is known to be the algebra of asymptotic symmetries at the null infinity of asymptotically-flat spacetimes in 2+1 dimensions [43].

Since these charges will become the basic operators in the quantum theory, it is important to understand their physical meaning (Sec. 11.1). In fact, as the Hilbert space is always assumed to be separable (i.e., to have a countable topological basis) in quantum mechanics, there is nothing distinguishing about the Hilbert space itself (other than its dimension) and all the information characterizing a particular theory lies in the way that relevant physical observables are represented on the Hilbert space (see footnote 5). The canonical charges derived from Isham's method have explicit formulas in term of variables describing the reduced phase space, but since these are quite abstract the underlying spacetime meaning of the charges must be uncovered. We explain (Sec. 11.1.1) that the P charges are associated with diffeomorphisms acting non-trivially at the boundary and their value are related to Fourier modes of $K_{ab}t^a s^b$, the components of the extrinsic curvature K_{ab} of the CMC slice along the tangent (t^a) and normal (s^a) unit vectors at the boundary of the disc. Of notable mention, the charge P_0 generates uniform rotation of the boundary (i.e., $SO(2)$ isometries of the boundary metric) and therefore can be interpreted as the spin (or angular momentum) of the diamond; moreover, its value is simply related to the twist (Sec. 11.1.2) of the corner loop.[7] The physical interpretation of the Q charges, on the other hand, has been far more elusive (Sec. 11.1.3). We know some properties that they must satisfy, and can speculate on what they could measure, but their precise meaning will be left for future examination.

Finally, we arrive at the quantum theory (Sec. 12), constructed from a (projective) unitary irreducible representation of canonical group $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$ or rather, as revealed from the Poisson algebra of canonical charges, the group BMS_3. Since this group is a semi-directed product of a group $G = \mathit{Vira}$ with an abelian group $\mathfrak{vira}$, Mackey's theory of induced representations [44] can be employed to obtain the unitary irreducible representations of $\widetilde{G}$ [45] (modulo possible limitations associated with the infinite dimensionality of the group [46]). The result (Sec. 12.1) is that the quantum theory is characterized by a choice of coadjoint orbit of Virasoro together with a choice of (projective) unitary irreducible representation of the corresponding little group (i.e., the subgroup of *Vira* that leaves any particular point of the orbit fixed). The most natural choice is to take the orbit diffeomorphic to the con-

[7]The *twist* of a loop embedded in a three-dimensional space is the integrated torsion (with respect to proper length) along the loop. The torsion measures how a normal frame that is carried along the curve, as close as possible to being parallel transported (more precisely, Fermi-Walker transported) rotates around the axis aligned with the curve. The twist can also be interpreted as measuring the holonomy for (Fermi-Walker) transporting a frame along the loop, i.e., after the completion of the loop the transported frame will return boosted with respect to the original frame, and the boost angle is equal to the twist.

figuration space, in which case the little group is $PSL(2,\mathbb{R}) \times \mathbb{R}$, where $\mathbb{R}$ comes from the central element of *Vira* and is assumed to be represented trivially. We justify this choice from the Casimir matching principle (Sec. 12.2), which proposes that the value of Casimir operators should be matched between the classical and the quantum.[8] In particular, we consider two types of Casimir operators related to the monodromy-class and winding number of the coadjoint orbits of Virasoro. With this refinement, the Hilbert space is realized (Sec. 12.3) by wavefunctions on $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ valued in unitary irreducible representations of $PSL(2,\mathbb{R})$. The quantum theory can also be directly constructed from the algebra, which is the quantization of the Poisson algebra above,

$$
\begin{aligned}
&[\widehat{P}_n, \widehat{P}_m] = (m-n)\widehat{P}_{n+m} \\
&[\widehat{Q}_n, \widehat{P}_m] = (m-n)\widehat{Q}_{n+m} + 4\pi n^3 \delta_{n+m,0} \\
&[\widehat{Q}_n, \widehat{Q}_m] = 0
\end{aligned}
\tag{1.4}
$$

but note that $\widehat{P}_n$ and $\widehat{Q}_n$ are not expected to be self-adjoint but rather to satisfy $(\widehat{P}_n)^\dagger = \widehat{P}_{-n}$ and $(\widehat{Q}_n)^\dagger = \widehat{Q}_{-n}$, mimicking the classical relations $(P_n)^* = P_{-n}$ and $(Q_n)^* = Q_{-n}$. We use this to derive the spectrum of $\widehat{P}_0$.

[8] Classically, Casimir operators Poisson-commute with a complete algebra of observables and therefore are constant functions on the phase space. Quantum-mechanically, Casimir operators commute with all other operators and therefore are realized as a multiple of the identity in any (complex) irreducible representation. At either level, the Casimir takes the same value on any physical state, so it is natural to assume that the value of a quantum Casimir should match with with the value of its classical counterpart. (This principle leads to Dirac's electric-magnetic charge quantization when quantizing a particle on a sphere from Isham's perspective [47].)

We discuss two important observables in the quantum theory. First, we show (Sec. 12.5) that the twist $\mathcal{T}$ of the corner loop, also related to the spin of the diamond, is quantized in the quantum theory according to

$$\mathcal{T} = \frac{16\pi^2 \ell_P}{\ell}(n+s)\,, \quad n \in \mathbb{Z} \tag{1.5}$$

where $s = 0$ or $1/2$ and ℓ is the total length of the corner loop and ℓ_P is the Planck length. Lastly, the charge $\widehat{Q}_0$, which happens to commute with the spin, can be shown (Sec. 12.4) to have a continuum spectrum that is bounded from above and unbounded from below, attaining a maximum value of 2π at a (generalized) state described by a wavefunction $\Psi([\psi])$ localized at the $PSL(2,\mathbb{R})$-class of the identity $[\psi] = [I]$.

It is worthwhile to notice that this accomplishes only the kinematical part of the quantization, i.e., characterizing the Hilbert space and the manner that the canonical observables are represented on it. A complete quantization would also involve the description of the dynamics, such as successfully representing the time-evolution Hamiltonian (discussed in Part I) and possibly other relevant observables which are expressible in terms of the canonical ones. The limitations and future prospects of our quantization are considered in the discussion (Sec. 13).

In addition to App. A — *glossary, symbols and conventions* — we include three other appendices. App. D describes the topology of the configuration space, $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. App. E offers a general review of Mackey's theory of induced representation and the associated notion of systems of imprimitivity. App. F discusses projective representations and its relationship to central extension (by 2-cocycles) of the group.

Part I

Classical reduction

2 Reduced phase space

One of the fundamental assumptions of physics is that the laws of nature should be deterministic. That is, the knowledge of the state of a (closed) system at a given time should allow us to completely know its state at any future (or past) time. More precisely, let the state of the system be denoted by ψ; if the system is at the state ψ_0 at the initial time $t = 0$, then the equations of motion should uniquely determine $\psi(t) = F(\psi_0; t)$ as a function F of the time t and the initial state. Certain theories, however, are described by equations of motion that are not deterministic, so that to each initial condition there corresponds a class of solutions $[\psi]$ to the equations of motion that satisfies the initial conditions. The principle of determinism then implies one of two things: (i) the theory is incomplete, so that further equations of motion or constraining conditions are necessary to make the time evolution unique; (ii) there is a redundancy in the description, so that only the equivalence classes of solutions, compatible with the initial conditions, are really physical. Thus, assuming that the theory is complete leave us with option (ii), meaning that the physical space of states is only a quotient of the prototypical space of states, under the projection $\psi \mapsto [\psi]$. Such theories are called *gauge theories*, and states ψ and ψ' within the same equivalence class, $[\psi] = [\psi']$, are said to be "gauge-equivalent" or "related by a gauge transformation".

A notorious example of a gauge theory is electromagnetism. The equations of motion are $d\,{}^\star F = 4\pi\,{}^\star J$, where $F := dA$ is the electromagnetic strength and A is the electromagnetic potential. Since the equations of motion depend only on F, it is insensitive to the change $A \mapsto A + d\alpha$, where α is any real function on the spacetime. Therefore, for any solution A of the equations of motion, and any function α that vanishes in a neighborhood of the spatial slice at $t = 0$, we have that $A' = A + d\alpha$ is also a solution of the equations of motion, satisfying the same initial conditions. In this way, the theory is incomplete unless we admit that the change $A \mapsto A + d\alpha$ is not physically observable, i.e., the physical states are described by the equivalence classes $[A] = [A + d\alpha]$.

The consequence of a gauge ambiguities for the Hamiltonian description is the appearance of constraints in the phase space [48]. To see this, let us consider a system described by a finite set of configuration variables q^i, $i = 1, \ldots n$, associated with an action principle $S = \int dt\, L(q, \dot{q})$. The equation of motion are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0 \tag{2.1}$$

From the chain rule it follows

$$\frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j = -\frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}\dot{q}^j + \frac{\partial L}{\partial q^i} \tag{2.2}$$

Note that a necessary and sufficient condition for this set of equations to have a unique solution, given initial values for q and $\dot{q}$, is that we can solve algebraically for the second time derivatives in terms of the lower time derivatives, i.e., that we can put it in the form $\ddot{q}^i = E^i(q, \dot{q})$. This is equivalent to say that

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\right) \neq 0 \tag{2.3}$$

where the argument of the determinant is understood as a $n \times n$ matrix with indices i and j. If we define the momenta p conjugated to q by

$$p_i := \frac{\partial L}{\partial \dot{q}^i} \tag{2.4}$$

then the non-degeneracy condition translates into

$$\det\left(\frac{\partial p_i}{\partial \dot{q}^j}\right) \neq 0 \tag{2.5}$$

This determinant is equal to the Jacobian of the Legendre transformation $(q, \dot{q}) \mapsto (q, p)$, mapping the space of "positions and velocities" to the phase space, and the non-degeneracy condition thus implies that this map is (locally) invertible. In a gauge system the equations of motion do not have unique solution, which requires (2.3) to be zero and therefore implies that the transformation $(q, \dot{q}) \mapsto (q, p)$ is not (locally) invertible. In other words, the image of the Legendre transformation is a surface $\mathcal{S}$ (of dimension less than $2n$) in the phase space $\mathcal{P}$. Hence there are constraints $C_\alpha = 0$ on the phase space specifying the location of this surface.

Let us assume that the constraint surface $\mathcal{S}$ is a manifold smoothly embedded in the phase space, and call $\rho : \mathcal{S} \to \mathcal{P}$ the embedding map. The symplectic form on $\mathcal{P}$, $\Omega = \delta p_i \wedge \delta q^i$, can be pulled back to the constraint surface, defining a "symplectic form" $\omega := \rho^* \Omega$ on $\mathcal{S}$. The reason for the quotations is that although ω is closed (since $\delta\omega = \delta(\rho^*\Omega) = \rho^*\delta\Omega = 0$) it may fail to be non-degenerate, i.e., there may exist certain *null* directions ξ along $\mathcal{S}$ such that $\imath_\xi \omega = 0$. If it happens that ω is degenerate, so it is called a pre-symplectic form, then the null directions are precisely the gauge directions. This is because given any Hamiltonian $H : \mathcal{S} \to \mathbb{R}$, the vector flow X generated by H is defined to be a solution of $\delta H = -\imath_X \omega$, but it is clear that for any solution X then $X' = X + \xi$ is also a solution. Therefore ξ are the directions corresponding to the ambiguities in the time evolution, and thus moving along ξ must not correspond to a change in the physical states. An interesting fact about the gauge directions is that they are surface-integrable, in the sense that the vector field commutator of two null fields ξ and ξ' is another null field. This follows from the identities $\imath_{[X,Y]} = [\pounds_X, \imath_Y]$ and $\pounds_X = \delta\imath_X + \imath_X \delta$, which applied to ω gives

$$\imath_{[\xi,\xi']}\omega = \pounds_\xi \imath_{\xi'}\omega - \imath_{\xi'}\left(\delta\imath_\xi\omega + \imath_\xi\delta\omega\right) = 0 \tag{2.6}$$

where it was used that $\imath_\xi\omega = \imath_{\xi'}\omega = 0$ and $\delta\omega = 0$. The set of all null vector fields thus form an (infinite-dimensional) algebra, and the set of all gauge transformations (flows along null directions) form a group. The "integrated" surfaces whose tangent vectors at every point are null are called *gauge orbits*. A simple linear algebra argument reveals that the dimension of such surfaces must be at most equal to the number of constraints (in fact, it matches the number of *first class* constraints[9]).

The *reduced phase space*, $\widetilde{\mathcal{P}}$, is the space of physically distinct states, defined by the quotient of $\mathcal{S}$ under the gauge transformations, or in other words, the equivalence classes of gauge orbits. There is a natural symplectic form $\widetilde{\omega}$ (closed and non-degenerate) on the reduced phase space, having the property that its pull-back under the quotient map $J : \mathcal{S} \to \widetilde{\mathcal{P}}$ is equal to ω, i.e., $\omega = J^*\widetilde{\omega}$. This can be seen by a constructive approach. Given $p \in \widetilde{\mathcal{P}}$, let $\bar{p} \in \mathcal{S}$ be a point in the pre-image of p under J, i.e., $J(\bar{p}) = p$. Given two vectors $X, Y \in T_p\widetilde{\mathcal{P}}$, consider any two vectors $\bar{X}, \bar{Y} \in T_{\bar{p}}\mathcal{S}$ that project to X and Y under J_*, i.e., $J_*\bar{X} = X$ and $J_*\bar{Y} = Y$. Define,

$$\widetilde{\omega}(X,Y)\big|_p := \omega(\bar{X},\bar{Y})\big|_{\bar{p}} \tag{2.7}$$

Now we must show that this definition is consistent, in the sense that it must not depend on the particular point $\bar{p}$ chosen in the pre-image of J, nor on the particular vectors chosen in the pre-image of J_*. Note that if ϕ is a gauge transformation on $\mathcal{S}$ implementing a flow along null vector fields, then by definition of the projection

[9] A constraint is *first-class* if its symplectic flow (under the original pre-symplectic form Ω) is tangent to the constraint surface, and it is *second-class* otherwise. A first-class constraint always leads to a degenerate ω: if C is a first-class constraint then its flow X, defined from $\delta C = -\imath_X\Omega$, will be tangent to $\mathcal{S}$; so if T is any vector tangent to $\mathcal{S}$ we have $0 = \delta C(T) = -\Omega(X,T) = -\omega(X,T)$, implying that X is a null direction for ω.

map we have $J \circ \phi = J$. This implies that the kernel of J_* at $\bar{p}$ consists of null vectors ξ at $\bar{p}$, i.e., $J_*\xi = 0$ if and only if $\imath_\xi \omega = 0$. Therefore $\bar{X}$ is defined up to the addition of a null vector, ξ. As replacing $\bar{X}$ by $\bar{X} + \xi$ in the right-hand side of (2.7) does not affect the result, this definition is insensitive to the choices of vectors in the pre-image of J_*. Now we consider the ambiguity in the choice of $\bar{p}$. First note that gauge transformations ϕ are symmetries of ω, $\phi^*\omega = \omega$, which follows from $\pounds_\xi \omega = \delta \imath_\xi \omega + \imath_\xi \delta\omega = 0$. Second note that for any other point $\bar{p}'$ in the pre-image of p under J, there must exist a gauge transformation ϕ such that $\bar{p}' = \phi(\bar{p})$. It follows from $J_*\phi_* = J_*$ that $\bar{X}' := \phi_*\bar{X}$ and $\bar{Y}' := \phi_*\bar{Y}$ are vectors at $\bar{p}'$ in the pre-image of X and Y under J_*. Hence,

$$\omega(\bar{X}', \bar{Y}')\big|_{\bar{p}'} = \omega(\phi_*\bar{X}, \phi_*\bar{Y})\big|_{\phi(\bar{p})} = \phi^*\omega(\bar{X}, \bar{Y})\big|_{\bar{p}} = \omega(\bar{X}, \bar{Y})\big|_{\bar{p}} \tag{2.8}$$

showing that (2.7) is independent of the choice of $\bar{p} \in J^{-1}(p)$. Therefore we have shown that there exists a antisymmetric 2-form $\widetilde{\omega}$ on $\widetilde{\mathcal{P}}$ satisfying $\omega = J^*\widetilde{\omega}$. To finish the story, we must show that $\widetilde{\omega}$ is closed and non-degenerate. Note that $0 = \delta\omega = J^*\delta\widetilde{\omega}$. But since J is a projection map, so J_* has maximum rank, then $J^*\delta\widetilde{\omega} = 0$ implies that $\delta\widetilde{\omega} = 0$, establishing closedness. Now let η be a vector on $\widetilde{\mathcal{P}}$ such that $\imath_\eta\widetilde{\omega} = 0$, and let $\bar{\eta}$ be a vector on $\mathcal{S}$ that projects to η under J, i.e., $J_*\bar{\eta} = \eta$. We have,

$$\imath_{\bar{\eta}}\omega(\cdot) = \omega(\bar{\eta}, \cdot) = J^*\widetilde{\omega}(\bar{\eta}, \cdot) = \widetilde{\omega}(J_*\bar{\eta}, J_*(\cdot)) = \imath_\eta\widetilde{\omega}(J_*(\cdot)) = 0 \tag{2.9}$$

implying that $\bar{\eta}$ is a null vector. But null vectors are in the kernel of J_* and hence $\eta = 0$, establishing non-degeneracy.

The definition of the reduced phase space above suggests a explicit procedure for its construction. The reduction process will generally go through the following steps:

1. From the Lagrangian L, compute the conjugate momenta, $p = \partial L/\partial \dot{q}$, and see whether there are constraints among the phase space variables, $C^1(q,p) = 0$. The superscript "1" is because these constraints coming directly from the Lagrangian, without imposing the equations of motion, are called *primary* constraints (not to be confused with "first-class" — footnote 9).

2. Compute the Hamiltonian H and find all *secondary* constraints C^2 coming from the equations of motion, i.e., following from imposing that the primary constraints must be respected by time evolution, $dC^1/dt = \{C^1, H\} \approx 0$, where the approximate sign indicates that the equality holds when all constraints are satisfied (note that the Poisson brackets are evaluate without imposing any constraints). The same process must be repeated for the secondary constraints, and tertiary constraints and so forth until no additional constraints are found. Once this process ends we can forget about this classification into primary, secondary, etc, and simply treat all the constraints on the same footing.

3. The set of all constraints determines the surface $\mathcal{S}$, so we compute the re-

striction of the original (pre)symplectic form Ω to $\mathcal{S}$, $\omega := \Omega|_{\mathcal{S}}$. If ω is non-degenerate, $\mathcal{S}$ is already the reduced phase space (this is the case if all constraints are second-class). If ω is degenerate, we must identify its null directions and the corresponding gauge orbits.

4. To find the reduced phase space, we mention two convenient ways to quotient by the gauge orbits that may be useful:

    - The first way is to "gauge-fix", which means introducing additional constraints in such a way that all constraints become second-class. In geometrical terms, we would look for a submanifold $\mathcal{S}'$ of $\mathcal{S}$ with the property that it intersects with each gauge orbit at precisely one point. In this case, the reduced phase space $\widetilde{\mathcal{P}}$ can be identified with $\mathcal{S}'$, and the symplectic form $\widetilde{\omega}$ is simply the restriction of ω to $\mathcal{S}'$. A gauge-fixing approach is not always possible, something known as the Gribov phenomenon. As an example, consider the case where the gauge orbits are isomorphic to a group, G, in such a way that $\mathcal{S}$ is a principal G-bundle over the reduced phase space $\widetilde{\mathcal{P}}$; in this case, the existence of a "gauge-fixing condition" corresponds to say that there exists a global cross section $\mathcal{S}'$ of $\mathcal{S}$, which is only possible if the bundle is trivial, i.e., $\mathcal{S} \sim \widetilde{\mathcal{P}} \times G$.

    - The second way is to "change coordinates" in a suitable manner. If it is possible to cover $\mathcal{S}$ with a coordinate system, there may be classes of coordinates such that the pre-symplectic form becomes simpler, and the null direction become evident. For example, consider $\mathcal{S} \sim \mathbb{R}^4$, covered with coordinates $\{x_1, x_2, p_1, p_2\}$ and a pre-symplectic form $\omega = \delta(p_1 + p_2) \wedge \delta(x_1 + x_2)$; this symplectic form suggests a change of coordinates to $x_\pm = x_1 \pm x_2$, $p_\pm = p_1 \pm p_2$, so that $\omega = \delta p_+ \wedge \delta x_+$, revealing that the appropriate quotient map is $J(x_1, x_2, p_1, p_2) = (x' := x_+, p' := p_+)$, where $\{x', p'\}$ are coordinates on $\widetilde{\mathcal{P}} \sim \mathbb{R}^2$. This approach may also not be always possible, as $\mathcal{S}$ may not admit a global coordinate system.

Note that both ways can be used in conjunction, each one removing part of the gauge ambiguities. Still these two procedures may not be enough, as there may remain some residual gauge ambiguities that have to be addressed in a specialized manner. When treating the causal diamond, we will employ the first way to deal with the ambiguities associated with the time diffeomorphisms, the second way to deal with the spatial diffeomorphisms, and there will remain a small (finite) number of ambiguities that will need to be removed in a third way.

To wrap up, let us return to the discussion at the beginning of this section about deterministic evolution. If the equations of motions are deterministic, then there is a one-to-one correspondence between solutions to the equations of motion and initial data. Since the phase space is the space of "initial data" (as it is isomorphic to the space of "initial positions and velocities"), we can look at the phase space as the space of "solutions to the equations of motion". This perspective is taken as the basis of the *covariant* construction of the phase space for field theories, which avoids having to pick an arbitrary time direction and spatial slice and simply consider the field configurations that satisfy the equations of motion in spacetime. In this construction, one starts with the space $\mathcal{C}$ of all field configurations on spacetime and define the phase space $\mathcal{P}$ as the submanifold consisting of configurations satisfying the equations of motion. The (pre)symplectic structure follows from the action, and it may be degenerate on $\mathcal{P}$. Accordingly, this degeneracy corresponds to gauge ambiguities and the reduced phase space is defined by quotienting over all the gauge transformations. In this way, we can think of the reduced phase space as the space of all *physical* solutions to the equations of motion (where gauge-equivalent solutions are regarded as the same physical solution).

Since the reduced phase space is a "standard" phase space, in the sense that it has a (non-degenerate) symplectic structure, it is amenable to the application of canonical quantization. In particular, there are no subtleties associated with gauge ambiguities when it comes to the quantization, for all the gauge has already been eliminated at the classical level. The fact that only the physical, gauge-invariant degrees of freedom are undergoing quantization is a highly appealing feature of this approach, giving some confidence that the class of quantum theories obtained are "natural". Nevertheless, this approach does not come without drawbacks. One of them is that the reduced phase space usually has a non-trivial topology, or at least does not have a natural vector space structure, which requires more sophisticated schemes of canonical quantization, such as geometric quantization or group-theoretic quantization. Also, certain observables like the time-evolution Hamiltonian may become complicated when written in terms of the "canonical observables" (i.e., the "q's and p's") compatible with the phase space topology, and this may lead to severe operator-ordering ambiguities. Since there is no general proof that the Dirac approach is equivalent to the the reduced phase space approach (in fact, there are examples where they are known to disagree — for a study in the case of general relativity see, e.g., [49]), it is worth to explore both approaches, and in this work we focus on the latter.

3 The causal diamond

In this section we define the dynamical system of interest, the causal diamond. We also provide a brief outline of the phase space reduction procedure, discussing the details in the subsequent sections.

3.1 The system

We define the causal diamond as the domain of dependence of a spatial slice Σ having the topology of a disc D (i.e., a 2-dimensional ball). More precisely, we define the spacetime as the maximal development of the initial-value problem associated with

Einstein's equations for pure gravity with a nonpositive cosmological constant Λ,

$$\mathcal{G}_{ab} + \Lambda g_{ab} = 0 \tag{3.1}$$

given a Riemannian metric h_{ab} and extrinsic curvature K^{ab} on $\Sigma \sim D$. Our choice of boundary conditions, to be later justified, is that the induced metric on the spatial boundary is fixed

$$h\big|_{\partial\Sigma} = \gamma \tag{3.2}$$

where γ is a given (fixed) metric on $\partial\Sigma \sim S^1$. Since we can parametrize the points of the boundary by the length with respect to γ, the only *intrinsic* attribute of this metric is the total length, ℓ, of the boundary.

According to the ADM (Arnowitt-Deser-Misner) formalism [50], the (pre)phase space $\mathcal{P}$ corresponds to the space of all Riemmanian metrics, satisfying the Dirichlet boundary condition, together with the space of all extrinsic curvatures on the ball $\Sigma \sim D$. For definiteness, let us denote the space of Riemannian metrics on the spatial disc by $\mathrm{Riem}(D;\gamma)$, consisting of all positive-definite symmetric tensors of type $({}^0{}_2)$, h_{ab}, on D satisfying $h|_{\partial D} = \gamma$; and denote the space of all extrinsic curvatures on D by $\mathrm{Sym}(D, ({}^2{}_0))$, consisting of all symmetric tensors of type $({}^2{}_0)$, K^{ab}, on D. The (pre)phase space then have the trivial product structure

$$\mathcal{P} = \mathrm{Riem}(D;\gamma) \times \mathrm{Sym}(D, ({}^2{}_0)) \tag{3.3}$$

One could worry about the degree of smoothness of these function spaces. In the study of partial differential equations, such as the initial value problem of Einstein's equation, it is natural to consider *Sobolev spaces*. The Sobolev space $W^{k,p}(U)$, where $k \in \mathbb{N}$, $1 \leq p \leq \infty$ and U is a open subset of $\mathbb{R}^n$, is defined as the space of all functions $f : U \to \mathbb{R}$ such that f and its (weak) derivatives of order equal or less than k are in $L^p(U)$. (The generalization for functions valued in $\mathbb{R}^n$ is natural.) These spaces are convenient for they are Banach spaces (complete normed vector spaces), which facilitates certain proofs of existence of solutions by allowing one to construct sequences of approximate solutions that converge to an exact solution. The case $p = 2$ is particularly interesting because it is a Hilbert space, implying that its (topological) dual is isomorphic to itself. This allows us to think of the phase space as a cotangent bundle, in a precise way, as we explain next. Note that $\mathrm{Riem}(D)$ can be seen as the open region of $\mathrm{Sym}(D, ({}^0{}_2))$ defined by the conditions $\det(h) > 0$ and $\mathrm{tr}(h) > 0$. Since $\mathrm{Sym}(D, ({}^0{}_2))$ is a linear space, we can assume that it is a Sobolev space $W^{k,2}(D)$ of symmetric-matrix-valued functions on D. Then the topology of $\mathrm{Riem}(D)$ is inherited from $\mathrm{Sym}(D, ({}^0{}_2))$. A tangent vector to $\mathrm{Riem}(D)$ can be seen as a vector in $\mathrm{Sym}(D, ({}^0{}_2))$, and due to the linearity of this space, the tangent vector can be naturally identified with an element of $\mathrm{Sym}(D, ({}^0{}_2))$ itself. We can write,

$$T_h[\mathrm{Riem}(D)] \sim \mathrm{Sym}(D, ({}^0{}_2)) \tag{3.4}$$

where T_h denotes the tangent space at $h \in \mathrm{Riem}(D)$. The space of 1-forms at h

can be defined to be the (topological) dual of $\mathrm{Sym}(D, ({}^0{}_2))$, and since that is a Hilbert space, the dual is isomorphic to itself. Nevertheless, it is most natural to characterize the space of 1-forms as $\mathrm{Sym}^1(D, ({}^2{}_0))$, the space of symmetric tensor densities of type $({}^2{}_0)$ and weight 1. In this way, the action of a dual vector π on a vector ζ, at h, is given by contracting them and integrating over D,

$$\pi(\zeta) := \int_D d^n x\, \pi^{ab} \zeta_{ab} \tag{3.5}$$

The space of tensor densities is isomorphic to the space of tensors since they can be related by a factor of a power of $\sqrt{\det(h)}$; in particular, $\pi^{ab} = \sqrt{\det(h)}\widetilde{\pi}^{ab}$, where $\widetilde{\pi}^{ab}$ is a standard tensor. Therefore,

$$T_h^*[\mathrm{Riem}(D)] \sim \mathrm{Sym}^1(D, ({}^2{}_0)) \sim \mathrm{Sym}(D, ({}^2{}_0)) \tag{3.6}$$

where T_h^* denotes the cotangent space at $h \in \mathrm{Riem}(D, \gamma)$. Since $\mathrm{Riem}(D)$ is an open subset in a vector space, its tangent and cotangent bundles are trivial. So,

$$T^*[\mathrm{Riem}(D)] = \mathrm{Riem}(D) \times \mathrm{Sym}(D, ({}^2{}_0)) \tag{3.7}$$

Our configuration space, however, contains the additional restriction on the induced boundary metric, so it is $\mathrm{Riem}(D, \gamma)$. This condition affects the tangent space, since tangent vectors ζ, tangent to curves $\mathrm{Riem}(D, \gamma)$, are now subjected to the homogeneous boundary condition

$$\zeta_{ab} t^a t^b \big|_{\partial D} = 0 \tag{3.8}$$

where t^a is a vector on D tangent to ∂D. Nonetheless, since the condition above refers to a set of measure zero in D, it does not affect the space of cotangent vectors, which can still be taken to be elements π of $\mathrm{Sym}^1(D, ({}^2{}_0))$ acting on vectors as in (3.5). The cotangent bundle thus continues to have a trivial structure,

$$T^*[\mathrm{Riem}(D,\gamma)] = \mathrm{Riem}(D,\gamma) \times \mathrm{Sym}(D, ({}^2{}_0)) \tag{3.9}$$

which equals the (pre)phase space $\mathcal{P}$ in (3.3). The symplectic structure will be described next.

3.2 An outline of the reduction process

Here we present a brief outline of the gauge reduction process for the diamond, based on the gauge-fixing of time by CMC slices, following a general program of quantization via phase space reduction introduced by Moncrief et al [9, 10, 51]. (See also a more recent paper reviving this program [39].) The details will be explained in the main sections of the paper.

Let us first discuss the constraints on the phase space. For this class of diamond-shaped spacetimes, it is natural to set up the ADM decomposition with respect to a family of surfaces anchored at the S^1 boundary. In fact, these are the only Cauchy surfaces in a diamond. This corresponds to restricting to lapse functions that vanish at the boundary, $N|_{\partial\Sigma} = 0$, and shift vectors that are tangent to the boundary, $N^a|_{\partial\Sigma} \in T(\partial\Sigma) \subset T(\Sigma)$. The action functional can be defined between any two such surfaces as

$$S = \frac{1}{16\pi G} \int d^3x \sqrt{-g}(\mathcal{R} - 2\Lambda) \tag{3.10}$$

up to boundary terms. (Notice that $\mathcal{R}$ stands for the Ricci scalar associated with the spacetime metric g, while the symbol R will be reserved for spatial metrics.) We wish to define the system in such a way that the spacetime solution (within the diamond) can be fully determined from initial data on a slice; we take the (pre)symplectic form Ω to be the conventional one from the ADM formalism[10]

$$\Omega = \int d^2x \, \delta\pi^{ab} \wedge \delta h_{ab} \tag{3.11}$$

[10] In field theories, the bulk part of the symplectic form is uniquely determined from the (bulk part of the) Lagrangian, but there are often ambiguities in choosing its boundary term [52, 53, 54]. In some settings, such as in dynamically-closed theories defined in a spacetime "cylinder" with appropriate conditions on the timelike boundary, one can (almost) determine the boundary term for the symplectic form by insisting that the action principle is well-posed (in the sense of admitting stationary points when the configuration variables are fixed at the initial and final Cauchy slices) [55]. As an open-system, it is unclear how to fix these ambiguities for a causal diamond without further structure; in particular, it seems that one needs to specify how the causal diamonds are to be embedded as subsystems of a larger spacetime, or describe condition for the symplectic flux across the horizons [56]. (See, however, [57].) Here we take the simplest choice, where the symplectic form is just the integral over the Cauchy slice of a local symplectic current, which is consistent with a self-contained causal diamond.

where π^{ab} is the momentum conjugate to the spatial metric h_{ab} defined by

$$\pi^{ab} = \sqrt{h}\left(K^{ab} - Kh^{ab}\right) \tag{3.12}$$

with $K = K^{ab}h_{ab}$ being the trace of the extrinsic curvature. Note that "δ" denotes the exterior derivative in phase space.

The ADM Hamiltonian takes the pure-constraint form

$$H = \frac{1}{16\pi G}\int_\Sigma d^2x\left[N\sqrt{h}\left(K^{ab}K_{ab} - K^2 - R + 2\Lambda\right) - 2N_b\nabla_a\pi^{ab}\right] \tag{3.13}$$

where R and ∇ are respectively the Ricci scalar and covariant derivative on Σ associated with h_{ab}. The derivative of a density is defined by converting it to a tensor, applying the derivative and then converting back to a density of the same weight; that is, here we have $\nabla_a\pi^{ab} := h^{1/2}\nabla_a(h^{-1/2}\pi^{ab})$. Note that K^{ab} is an implicit function of h_{ab} and π^{ab} obtained by inverting (3.12); in 2+1 dimensions it is given by $K^{ab} = h^{-1/2}(\pi^{ab} - \pi h^{ab})$, where $\pi := \pi^{ab}h_{ab}$ is the trace of π^{ab}. The variation with respect to N yields the *Hamiltonian constraint*

$$K^{ab}K_{ab} - K^2 - R + 2\Lambda = 0 \tag{3.14}$$

and the variation with respect to N^a yields the *momentum constraint*

$$\nabla_a(K^{ab} - Kh^{ab}) = 0 \tag{3.15}$$

While the momentum constraint is linear in the momentum variables, similarly to electromagnetism, the Hamiltonian constraint is non-linear. This non-linearity is partially responsible for the relative difficulty in dealing with the constraints of gravity.

An interesting method for solving the constraints of general relativity is the *Lichnerowicz method* [58, 59, 60, 16, 9]. The goal is to convert the Hamiltonian constraint, which is a non-linear differential equation involving tensor fields K^{ab} and h_{ab}, into a differential equation for a single scalar field ϕ, which is accomplished by a suitable Weyl transformation. Let us introduce the *traceless part of the extrinsic curvature*, σ^{ab}, which in 2+1 dimension is given by

$$\sigma^{ab} := K^{ab} - \frac{1}{2} K h^{ab} \tag{3.16}$$

The constraints then become

$$\sigma^{ab}\sigma_{ab} - R + 2\Lambda - \frac{1}{2}K^2 = 0 \tag{3.17}$$

$$\nabla_a \left(\sigma^{ab} - \frac{1}{2} K h^{ab} \right) = 0 \tag{3.18}$$

Now suppose that we have a set of *input data*, (h_{ab}, σ^{ab}, K), that satisfies the momentum constraint but not necessarily the Hamiltonian constraint. The idea is to deform this data by a suitable pointwise conformal transformation (i.e., a Weyl transformation), preserving the momentum constraint, until the Hamiltonian constraint is satisfied. Consider the transformed data $(\widetilde{h}_{ab}, \widetilde{\sigma}^{ab}, \widetilde{K})$ defined by

$$\begin{aligned}
\widetilde{h}_{ab} &= e^{\phi} h_{ab} \\
\widetilde{\sigma}^{ab} &= f(\phi)\sigma^{ab} \\
\widetilde{K} &= w(\phi)K
\end{aligned} \tag{3.19}$$

where $\phi : \Sigma \to \mathbb{R}$ is a scalar function on Σ and $f : \mathbb{R} \to \mathbb{R}$ and $w : \mathbb{R} \to \mathbb{R}$ are real functions to be specified. (Note that $f(\phi)$ and $w(\phi)$ are to be understood as $f \circ \phi$ and $w \circ \phi$.) If we assume that the metric h_{ab} in the input data satisfies the boundary condition, we can impose the Dirichlet boundary condition on ϕ,

$$\phi\big|_{\partial\Sigma} = 0 \tag{3.20}$$

so that the deformed data $\widetilde{h}_{ab}$ also satisfies the same boundary condition for the induced metric. This will ensure that the physical initial data generated by this method (i.e., the data that satisfies both the momentum and the Hamiltonian constraints) will also satisfy the boundary conditions. The momentum constraint for the deformed data becomes,

$$\begin{aligned}
&\widetilde{\nabla}_a \left(\widetilde{\sigma}^{ab} - \frac{1}{2}\widetilde{K}\widetilde{h}^{ab} \right) = \\
&\quad = \frac{1}{2}\left(f(\phi) + w(\phi)e^{-\phi} \right) \nabla^b K + \left(f'(\phi) + 2f(\phi) \right) \sigma^{ab}\nabla_a\phi + \frac{1}{2}w'(\phi)e^{-\phi}K\nabla^b\phi = 0
\end{aligned} \tag{3.21}$$

where it was used that (h_{ab}, σ^{ab}, K) satisfies the momentum constraint, i.e., $\nabla_a \sigma^{ab} = \frac{1}{2}\nabla^b K$. Here $\widetilde{\nabla}$ is the covariant derivative associated with $\widetilde{h}_{ab}$. The Hamiltonian constraint becomes,

$$\widetilde{\sigma}^{ab}\widetilde{\sigma}_{ab} - \widetilde{R} + 2\Lambda - \frac{1}{2}\widetilde{K}^2 = \\ = e^{-\phi}\nabla^2\phi + e^{2\phi} f(\phi)^2 \sigma^{ab}\sigma_{ab} - e^{-\phi}R + 2\Lambda - \frac{1}{2}w(\phi)^2 K^2 = 0 \qquad (3.22)$$

In the left hand side, $\widetilde{\sigma}^{ab}\widetilde{\sigma}_{ab} = \widetilde{h}_{ac}\widetilde{h}_{bd}\widetilde{\sigma}^{ab}\widetilde{\sigma}^{cd}$; in the right hand side, $\sigma^{ab}\sigma_{ab} = h_{ac}h_{bd}\sigma^{ab}\sigma^{cd}$ and $\nabla^2 = h^{ab}\nabla_a\nabla_b$.

Given functions f and w, the constraints have thus become a pair of coupled differential equations for the scalar ϕ. For this method to be useful, we would like to be able to choose f and w in such a way that the equations become as simple as possible and, most importantly, that the equations admit solutions for a large class of input data (h_{ab}, σ^{ab}, K). The most useful choice comes up in conjunction with a gauge-fixing for the time: we consider the *constant mean curvature* gauge (abbreviated as "CMC"), in which the spatial slices are taken to have a constant trace of extrinsic curvature,

$$K\big|_{\Sigma} = -\tau \qquad (3.23)$$

where τ is a constant parameter on Σ. In the case of the diamond, we will establish that by varying τ from $-\infty$ to ∞ the spacetime will be foliated with Cauchy slices. Note that the "initial time", $\tau = 0$, corresponds to the maximal slice (i.e., the slice with maximal area). In this gauge, the first term of the momentum constraint (3.21) vanishes, suggesting the following convenient choices for the functions f and w,

$$
\begin{aligned}
f'(\phi) + 2f(\phi) = 0 \quad &\rightarrow \quad f(\phi) = e^{-2\phi} \\
w'(\phi) = 0 \quad &\rightarrow \quad w(\phi) = 1
\end{aligned}
\tag{3.24}
$$

With those choices the momentum constraint is automatically satisfied for any ϕ, provided that the input data satisfies

$$
\nabla_a \sigma^{ab} = 0 \tag{3.25}
$$

Also, note that ϕ is determined as the solution of a single differential equation coming from the Hamiltonian constraint (3.22),

$$
\nabla^2 \phi - R + e^{-\phi} \sigma^{ab} \sigma_{ab} - e^{\phi} \chi = 0 \tag{3.26}
$$

where

$$
\chi = -2\Lambda + \frac{1}{2}\tau^2 \tag{3.27}
$$

is a time-dependent parameter. Equation (3.26) for ϕ is called the *Lichnerowicz equation*. Note that $\Lambda \leq 0$ implies that $\chi \geq 0$ for all times, and this will be used for proving existence and uniqueness of solutions ϕ to the Lichnerowicz equation (for any Dirichlet boundary condition, such as in (3.20)). This means that for any input data $(h_{ab}, \sigma^{ab}, -\tau)$, in the CMC gauge, we can always generate a unique set of valid initial data given by $(e^{\phi} h_{ab}, e^{-2\phi}\sigma^{ab}, -\tau)$.

The Lichnerowicz method thus yields the following characterization for the constrained phase space, $\mathcal{S}$. As mentioned above, a set of input data $\mathscr{S} = (h_{ab}, \sigma^{ab}, -\tau)$ corresponds to one and only one set of valid initial data, $(e^{\phi} h_{ab}, e^{-2\phi}\sigma^{ab}, -\tau)$, where ϕ is the unique solution of (3.26). Now given any scalar field $\lambda : \Sigma \to \mathbb{R}$ vanishing at the boundary, $\lambda|_{\partial\Sigma} = 0$, consider a deformed set of input data $\mathscr{S}' = (e^{\lambda} h_{ab}, e^{-2\lambda}\sigma^{ab}, -\tau)$. The Lichnerowicz problem for $\mathscr{S}'$ leads to a unique valid data $(e^{\phi'} e^{\lambda} h_{ab}, e^{-2\phi'} e^{-2\lambda}\sigma^{ab}, -\tau)$, where ϕ' is the unique solution of

$$\nabla'^2 \phi' - R' + e^{-\phi'}(e^{-2\lambda}\sigma^{ab})(e^{-2\lambda}\sigma^{cd})(e^{\lambda}h_{ac})(e^{\lambda}h_{bd}) - e^{\phi'}\chi = 0 \tag{3.28}$$

with boundary condition $\phi'|_{\partial\Sigma} = 0$. Here ∇' and R' are associated with $h'_{ab} := e^{\lambda} h_{ab}$. This equation can be rewritten in terms of ∇^2 and R as,

$$e^{-\lambda}\nabla^2\phi' - e^{-\lambda}\left(R - \nabla^2\lambda\right) + e^{-\phi'} e^{-2\lambda}\sigma^{ab}\sigma_{ab} - e^{\phi'}\chi = 0 \tag{3.29}$$

or equivalently,

$$\nabla^2(\phi' + \lambda) - R + e^{-(\phi'+\lambda)}\sigma^{ab}\sigma_{ab} - e^{(\phi'+\lambda)}\chi = 0 \tag{3.30}$$

which has the same form as (3.26), with the same boundary condition, but in the variable $\phi' + \lambda$. Therefore $\phi' + \lambda = \phi$ is the unique solution, implying that the initial data obtained from $\mathscr{S}'$ is the same one obtained from $\mathscr{S}$, i.e., $(e^{\phi'} e^{\lambda} h_{ab}, e^{-2\phi'} e^{-2\lambda} \sigma^{ab}, -\tau) = (e^{\phi} h_{ab}, e^{-2\phi} \sigma^{ab}, -\tau)$. Consequently, each Weyl-related equivalence class of input data,

$$(h_{ab}, \sigma^{ab}, -\tau) \sim (e^{\lambda} h_{ab}, e^{-2\lambda} \sigma^{ab}, -\tau)\,, \quad \lambda\big|_{\partial\Sigma} = 0 \tag{3.31}$$

corresponds to a unique set of valid data $(e^{\phi} h_{ab}, e^{-2\phi} \sigma^{ab}, -\tau)$. In other words, the constraint surface $\mathcal{S}$ can be identified with this space of equivalence classes

$$\mathcal{S} = [(h_{ab}, \sigma^{ab}) \sim (e^{\lambda} h_{ab}, e^{-2\lambda} \sigma^{ab})] \tag{3.32}$$

where τ has been omitted since it does not transform under this Weyl transformation.

The constraint surface $\mathcal{S}$ contains null directions, corresponding to the gauge transformations associated with spatial diffeomorphisms on the CMC slices Σ. Let $\Psi : \Sigma \to \Sigma$ be a *boundary-trivial* diffeomorphism of Σ, i.e., a diffeomorphism that acts trivially on the boundary,

$$\Psi\big|_{\partial\Sigma} = \mathrm{I} \tag{3.33}$$

where I is the identity map. It is clear that if we apply such a diffeomorphism to a set of initial data we will produce a physically equivalent set of data.[11] In this

[11] Note that diffeomorphisms which are not boundary-trivial are not allowed for they generally do not preserve the Dirichlet boundary condition for the induced metric. The only exception is the $SO(2)$ family of isometries of the boundary, but those are true symmetries of the system, not gauge transformations.

way, the reduced phase space $\widetilde{\mathcal{P}}$ can be identified with the quotient of $\mathcal{S}$ under those boundary-trivial spatial diffeomorphisms. Using the above characterization for the constraint surface, (3.32), we can identify the reduced phase space as

$$\widetilde{\mathcal{P}} = \{(h_{ab}, \sigma^{ab}) \sim (\Psi_* e^{\lambda} h_{ab}, \Psi_* e^{-2\lambda} \sigma^{ab})\} \tag{3.34}$$

Note that the transformation of the metric, $h_{ab} \mapsto \Psi_* e^{\lambda} h_{ab}$, is a general boundary-trivial *conformal transformation*, i.e., a combination of a Weyl transformation and a diffeomorphism, (Ψ, λ), satisfying the boundary condition $(\Psi, \lambda)|_{\partial\Sigma} = (\mathrm{I}, 0)$.[12] We are going to review how the space of equivalence classes in (3.34) can be identified with the cotangent bundle of the space of conformal geometries on Σ. The space of conformal geometries on Σ, denoted by $\mathrm{ConGeo}(\Sigma)$, is the space of equivalence classes of Riemannian metrics quotiented by boundary-trivial conformal transformations,

$$\mathrm{ConGeo}(\Sigma) = \{h_{ab} \sim \Psi_* e^{\lambda} h_{ab}\} \tag{3.35}$$

Thus the phase space can be identified as

$$\widetilde{\mathcal{P}} = T^*[\mathrm{ConGeo}(\Sigma)] \tag{3.36}$$

This result quite general and comes up whenever the system admits a CMC gauge

[12]The nomenclature here may differ from other references, which sometimes restrict the term "conformal tranformation" to those that satisfy $h_{ab} = \Psi_* e^{\lambda} h_{ab}$; instead we refer to those very special conformal transformations as *conformal isometries*.

and the Lichnerowicz method is "nicely" posed (i.e., there are existence and uniqueness theorems for the associated Lichnerowicz equation). In the case of the diamond, we will see that the reduced phase space is given by

$$\widetilde{\mathcal{P}} = T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})] \tag{3.37}$$

where $\mathit{Diff}^+(S^1)$ is the group of orientation-preserving diffeomorphisms of the boundary $\partial\Sigma \sim S^1$ and $PSL(2,\mathbb{R})$ is the projective special linear group in 2 real dimensions (which is a 3-dimensional closed subgroup of $\mathit{Diff}^+(S^1)$, in a way that will be described later).

We will also discuss another approach for reducing the phase space, in Sec. 6. The idea is to "change coordinates" from (h_{ab}, σ^{ab}) to a new set of variables $(\Psi, \Phi, \sigma^{ab})$, where Ψ is a diffeomorphism of $\Sigma \sim D$ and $\Phi : \Sigma \to \mathbb{R}$ is a scalar function. This change of coordinates is defined by taking a standard metric $\bar{h}_{ab}$, such as the Euclidean metric on Σ (i.e., the metric of a flat disc, $dr^2 + r^2 d\theta^2$ with unit radius), and considering the conformal transformation (Ψ, Φ) from $\bar{h}_{ab}$ into $h_{ab} = \Psi_* \Phi \bar{h}_{ab}$. This is allowed because of the uniformization theorem, which ensures that any two metrics on a topological disc can be related by a conformal transformation. The contraint surface $\mathcal{S}$ can then be covered with "coordinates" $(\Psi, \bar{\sigma}^{ab})$, where $\bar{\sigma}^{ab}$ satisfies the divergenceless condition $\nabla_a \bar{\sigma}^{ab} = 0$ (where here ∇ is the derivative associated with $\bar{h}_{ab}$). The symplectic form becomes evidently degenerate with respect to "bulk diffeomorphisms", suggesting a reduction to a (not fully) reduced phase space that can be covered with "coordinates" $(\psi, \bar{\sigma}^{ab})$, where $\psi := \Psi|_{\partial\Sigma}$ is the boundary action of Ψ. This space has the topology $\mathit{Diff}^{+}(S^1) \times$ *"functions on S^1 not containing the Fourier modes 1, $\sin\theta$ and $\cos\theta$"*. By a simple analysis, we can determine that there are still three degenerate directions to be removed, corresponding to the $PSL(2,\mathbb{R})$ subgroup of $\mathit{Diff}^{+}(S^1)$, ultimately leading to the (fully) reduced phase space (3.37). Not only this alternative approach serves as a check of the earlier result, but it is also very useful as it provides an explicit "change of coordinates" from the "natural" variables describing the reduced phase space to the more easily interpretable geometric quantities like spatial metric and extrinsic curvature.

Before proceeding with more technical developments, it is interesting to understand what the classical states described in (3.37) actually represent. First note that in three spacetime dimensions there are no "local" gravitational degrees of freedom (i.e., there are no gravitational waves). This is because the Weyl tensor vanishes identically, meaning that the curvature is completely determined by the Ricci tensor,

$$\mathcal{R}_{abcd} = 2\left(g_{a[c}\mathcal{R}_{d]b} - g_{b[c}\mathcal{R}_{d]a}\right) - \mathcal{R}g_{a[c}g_{d]b} \tag{3.38}$$

and the Ricci tensor is fixed by Einstein's equation. Therefore the metric can be "locally" determined using normal coordinates,[13] leaving no physical (gauge-invariant) degree of freedom left. If there is no matter present, Einstein's equation imply that $\mathcal{R}_{ab} = 2\Lambda g_{ab}$, so

$$\mathcal{R}_{abcd} = 2\Lambda g_{a[c}g_{d]b} \tag{3.39}$$

implying that the metric is maximally symmetric. In the case $\Lambda < 0$ this means that the spacetime is locally Anti-de Sitter (AdS_3), and in the case $\Lambda = 0$ the spacetime is Minkowski ($Mink_3$). Thus our diamond is "locally AdS_3" (resp. "locally $Mink_3$), meaning that the neighborhood of any point in the bulk can be isometrically embedded into global AdS_3 (resp. $Mink_3$) spacetime. In fact, since we are considering trivial topology for the Cauchy slices, the entire diamond can be isometrically embedded as a region of AdS_3 (resp. $Mink_3$) spacetime. Consequently, the only degrees of freedom are associated with the global shape of the diamond. In other words, the space (3.37) must be describing diamond-shaped regions of AdS spacetime, with a fixed boundary length ℓ. It is reasonable to ask whether the phase space can be identified with a special subset (or classes of equivalence) of embedded diamonds in AdS_3. However, the formal analysis of the reduced phase space paired with the explicit embedding construction (App. C) there is no natural one-to-one

[13]The *Riemann normal coordinates* are defined with respect to the exponential map (associated with the metric) in the following way. Given a manifold $\mathcal{M}$ with metric g_{ab}, let $U \subset \mathcal{M}$ be a neighborhood of $x_0 \in \mathcal{M}$ such that the exponential map (based at x_0), $\exp_{x_0} : T_{x_0}\mathcal{M} \to \mathcal{M}$, is an isomorphism between an open neighborhood V of $T_{x_0}\mathcal{M}$ and U. This means that for any $x \in U$ there exists a unique $v \in V \subset T_{x_0}\mathcal{M}$ such that $x = \exp_{x_0}(v)$. Given a basis e_μ of $T_{x_0}\mathcal{M}$ we can decompose $v = x^\mu e_\mu$. The normal coordinates on U, with respect to x_0 and e_μ, is defined by assigning coordinates x^μ to x. The metric g_{ab} on U (or a open subset of U) can be reconstructed from the value of the curvature (and all its derivatives) at x_0, and it is given (to first order) by $g_{\mu\nu}(x) = \eta_{\mu\nu} - \frac{1}{3}R_{\mu\alpha\nu\beta}(0)x^\alpha x^\beta + \cdots$, where e_μ is taken to be orthogonal.

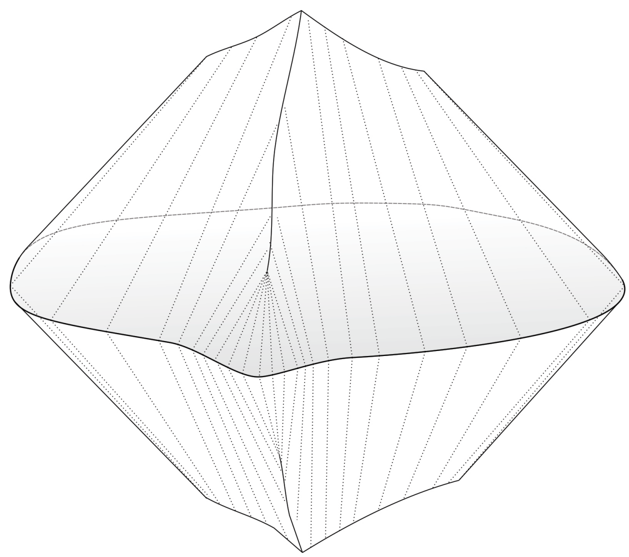


Figure 3: The phase space consists of causal diamonds in AdS_3 (or $Mink_3$ if $\Lambda = 0$) with topologically trivial Cauchy slices (discs) whose corner loops have fixed length ℓ.

correspondence between points in the phase space with any special subset of embeddings. We will explain this point in more detail in App. C. Finally, note that any such a diamond-shaped region in AdS_3 can be fully specified by giving how the S^1 boundary, with length ℓ, embeds into AdS_3. More precisely, let $C \subset AdS_3$ be a (spacelike, achronal) loop in AdS_3, with length ℓ, satisfying the condition that it is the boundary of a spacelike topological disc D, then the domain of dependence of D determines a unique diamond; moreover, observe that the diamond so defined depends only on C, but not on D.

4 Constant mean curvature foliation

In this section we show that the class of spacetimes consisting of causal diamonds, with nonpositive cosmological constant, admits a nice foliation by surfaces of constant mean curvature. Some pertinent references are [61, 62, 63, 64]. Moreover, the Lichnerowicz problem set up with respect to such a foliation is well-posed in the sense that the solutions to the Lichnerowicz equation exist and are unique given the boundary conditions.

4.1 The foliation is nicely behaved

One fundamental step in the reduction process is the preliminary gauge-fixing of time by the choice of a constant-mean-curvature (CMC) foliation of the spacetime. It is therefore essential that we can guarantee the a priori existence and regularity of such a foliation for all possible initial data, that is, all causal diamonds with fixed boundary metric. For motivation, we shall begin this section with a very simple argument based on Raychaudhuri's equation establishing some nice properties of CMC slices in our class of spacetimes, such as the fact that there can exist at most one CMC slice with a given $\tau = -\kappa$, and if two CMC slices exist such that $\tau_2 > \tau_1$ then CMC_2 is entirely to the future of CMC_1. In the next subsection, we cite a general theorem ensuring the existence and regularity of a foliation by CMCs. Finally we prove that a slice approaching the future horizon of the diamond has arbitrarily negative supremum of mean curvature ($\sup \kappa \to -\infty$) while a slice approaching the past horizon has arbitrarily positive infimum of mean curvature ($\inf \kappa \to \infty$), which implies that the foliation covers the whole causal diamond as τ ranges from $-\infty$ to ∞.

The Raychaudhuri's equation governs how a congruence of geodesic expands, twists and shears. If ξ the unit vector tangent to a congruence of timelike geodesics in a (1+d dimensional) spacetime with metric g_{ab}, and $h_{ab} := g_{ab} + \xi_a \xi_b$ is the "spatial metric" (i.e., $h^a_{\ b}$ is the projector onto the subspace orthogonal to ξ), we define the following parameters associated to the congruence: expansion $\theta := \nabla_a \xi^a$, shear $\sigma_{ab} := \nabla_{(a}\xi_{b)} - \frac{1}{d}\theta h_{ab}$ and twist $\omega_{ab} := \nabla_{[a}\xi_{b]}$. The equation describing how the geodesics expand in time is

$$\frac{d\theta}{ds} = \xi^a \nabla_a \theta = -\frac{1}{d}\theta^2 - \mathcal{R}_{ab}\xi^a \xi^b - \sigma_{ab}\sigma^{ab} + \omega_{ab}\omega^{ab} \tag{4.1}$$

where s is the proper length along the geodesics and $\mathcal{R}_{ab}$ is the Ricci curvature associated with g_{ab}. The equation for the twist is

$$\xi^c \nabla_c \omega_{ab} = -\frac{2}{d}\theta \omega_{ab} + 2\sigma^c_{\ [a}\omega_{b]c} \tag{4.2}$$

and we can see that if $\omega_{ab} = 0$ at one point of a geodesic then it will remain zero along that geodesic. Frobenius theorem says that the congruence is (locally) hypersurface orthogonal if and only if $\omega_{ab} = 0$; hence, if the congruence is defined by shooting geodesics orthogonally from a codimension-1 surface, then $\omega_{ab} = 0$.

Let Σ_1 and Σ_2 be two compact, acausal surfaces, sharing the same boundary, with constant mean curvatures κ_1 and κ_2, respectively. Take $\Gamma = \Sigma_1 \cap \Sigma_2$ as the set of points where the two surfaces intersect each other; Γ will divide Σ_1 and Σ_2 into patches, Σ_1^i and Σ_2^i, such that the spacetime region between each Σ_1^i and Σ_2^i is "lens shaped". More precisely, let us define Σ_1^i and Σ_2^i to be coverings of Σ_1 and Σ_2, respectively, by compact connected regions satisfying the following properties:

(i) Either $\Sigma_1^i \cap \Gamma = \partial\Sigma_1^i = \partial\Sigma_2^i = \Sigma_2^i \cap \Gamma$ or $\Sigma_1^i = \Sigma_2^i \subset \Gamma$;

(ii) If $\Sigma_1^i \neq \Sigma_2^i$, then $\mathrm{Int}(\Sigma_1^i)$ is either entirely to the future or entirely to the past of $\mathrm{Int}(\Sigma_2^i)$.

Now the argument can be made for each i. Let us first consider the case where $\kappa_1 \neq \kappa_2$, so $\Sigma_1^i \neq \Sigma_2^i$. Suppose that Σ_2^i is to the future of Σ_1^i and consider the set of all timelike geodesics from Σ_1^i to Σ_2^i; let γ be one geodesic with maximum length. Evidently the length of γ is non-zero. Moreover γ would be orthogonal to both Σ_1^i and Σ_2^i. Now consider a congruence of geodesics around γ starting orthogonal to Σ_1^i. Since this congruence starts non-twisting ($\omega_{ab} = 0$) it would remain hypersurface orthogonal; to a first-order approximation, Σ_2^i would be orthogonal to the congruence at the point where it intersects with γ. Therefore, at the point where Σ_1^i intersects with γ we have $\theta = \kappa_1$ and at Σ_2^i we have $\theta = \kappa_2$; also, σ_{ab} would correspond to the traceless part of the extrinsic curvature of Σ_1^i and Σ_2^i at the respective points. We can then integrate the Raychaudhuri's equation from Σ_1^i to Σ_2^i to get

$$\kappa_2 - \kappa_1 = \int ds\, \frac{d\theta}{ds} = -\int ds \left(\frac{1}{2}\theta^2 + \sigma_{ab}\sigma^{ab} - 2\Lambda \right) \tag{4.3}$$

where we have particularized to $d = 2$, used Einstein's equation $\mathcal{R}_{ab} = 2\Lambda g_{ab}$ and $\xi_a \xi^a = -1$. Note that, for a non-positive cosmological constant, the left-hand side is non-positive. Thus $\kappa_2 \leq \kappa_1$. That is, if a CMC is entirely to the future of another CMC, then the latter must not have a smaller mean curvature. We can easily strengthen the conclusion if we assume that (i) κ_1 and κ_2 are not both zero or (ii) the spacetime is negatively curved, $\Lambda < 0$. In case (i) we see that θ^2 must be non-vanishing on some portion of the maximal curve, and in case (ii) we have $-2\Lambda > 0$. In both cases the left-hand side of the equation is negative, implying that $\kappa_2 < \kappa_1$. In particular, this means that two CMCs sharing the same boundary, with different κ's, satisfying one of these conditions must not intersect at points in their interior; and the CMC with the smallest κ will be to the future of the other one. Note also that if the two κ's are infinitesimally close to each other then the CMCs must be infinitesimally close to each other, as can be seen from the inequality

$$\text{length}(\gamma) \leq \frac{|\kappa_2 - \kappa_1|}{\min_\gamma \left(\theta^2/2 + \sigma_{ab}\sigma^{ab}\right) - 2\Lambda} \tag{4.4}$$

We have therefore seen that, if one of these conditions are satisfied, then the CMCs (if they exist) would never intersect each other (except at their common boundaries), they would be temporally ordered (i.e., as κ decreases the CMCs move to the future) and they must foliate a region of the space (i.e., there can be no gap between CMCs with infinitesimally close κ's).

Not covered in the previous argument is the case of maximal slices ($\kappa = 0$) in flat spacetime ($\Lambda = 0$) with zero extrinsic curvature. The previous argument cannot not rule out the possibility that there are more than one maximal slice, and perhaps with a gap between them (i.e., a region between two maximal slices devoid of any CMCs). This simplest manner to approach this case is by considering a continuity argument. Namely, one can see that the foliation varies smoothly (in a given background manifold) with respect to infinitesimal variation of the Λ parameter; as the foliation is well-behaved for all $\Lambda < 0$, with no gap at $\kappa = 0$, which remains true in the limit $\Lambda \to 0^-$, implies that the foliation is also well-behaved at $\Lambda = 0$.

4.1.1 Crushing singularity

Consider a globally-hyperbolic connected spacetime, having a compact Cauchy slice Σ. All Cauchy slices are homeomorphic, so consider an arbitrary homeomorphism between any slice Σ to a reference slice Σ_0, which allows us to compare points x in different slices. Let Σ_- and Σ_+ be two (sufficiently regular) Cauchy surfaces (closed or with a common boundary) with mean curvatures $K_-(x)$ and $K_+(x)$, respectively, and suppose that Σ^+ is entirely to the future of Σ^-. Now, if $K_+(x) < K_-(x)$ then for any (continuous) function $K(x)$ such that $K_+(x) < K(x) < K_-(x)$, there exists

[64] a slice Σ, between Σ_- and Σ_+, whose mean curvature is $K(x)$.

Combined with the previous arguments establishing the nice properties of CMC slices, it follows the CMC foliation exists and spans the whole spacetime if one can show that the future and past horizons are *crushing singularities* [64, 65]. More precisely, the future horizon $\mathcal{N}^+$ (i.e., the boundary of the future domain of dependence of a Cauchy slice) is said to be a crushing singularity if there exists a family of surfaces Σ^+_λ such that $\lim_{\lambda\to\infty}\Sigma^+_\lambda = \mathcal{N}^+$ and $\lim_{\lambda\to\infty}\sup_{\Sigma^+_\lambda} K = -\infty$; and the past horizon $\mathcal{N}^-$ is similarly said to be a crushing singularity if there exists a family of surfaces Σ^-_λ such that $\lim_{\lambda\to\infty}\Sigma^-_\lambda = \mathcal{N}^-$ and $\lim_{\lambda\to\infty}\inf_{\Sigma^-_\lambda} K = +\infty$. In words, this means that one could find a family of surfaces approaching the future horizon whose mean curvature is arbitrarily negative at every point, and similarly a family a surfaces approaching the past horizon whose mean curvature is arbitrarily positive everywhere. Then for any constant $K \in (-\infty, +\infty)$, there will exist a slice with $K(x) = K$; as shown before, these slices will be unique (for each K) and continuously ordered in time, $\tau = -K$, with no gap (i.e., an open region not sliced by any CMC), thereby defining a regular CMC foliation of the whole diamond.

Here we shall argue that for any causal diamond in our phase space the future horizon is a crushing singularity. A completely analogous analysis can be used to show that the past horizon is also a crushing singularity. The proof will not be fully rigorous, but we hope that it will be sufficiently convincing. The idea is to consider a surface $\mathcal{S}$ that is very close to a null surface $\mathcal{N}$ (whose null generators are geodesic), and in a suitable coordinate system adapted to $\mathcal{N}$, argue using a first order approximation (in the parameter describing the nearness of $\mathcal{S}$ to $\mathcal{N}$) that we can define S with arbitrarily negative K. Note that $\mathcal{S}$ is not an entire Cauchy slice since $\mathcal{N}$ typically corresponds only to a portion of the future horizon, which is not a manifold because of the caustics. In fact, it appears that, in general, the future horizon can always be described by a finite number of null manifolds $\mathcal{N}_i$, emanating from the corner $\partial\Sigma$, and meeting at the graph-like caustics—see Fig. 12 for a representation of a typical shape of the horizon. We will consider a set of $\mathcal{S}_i$, near their respective $\mathcal{N}_i$, and join them smoothly in a neighborhood of the caustics, by "rounding off" their intersection.

First let us review the coordinate formula for the mean curvature of a surface. Suppose that in some open spacetime region, there are coordinates t and x^α such that the surface $\mathcal{S}$ can be described as the zero-set of the function

$$S := t - f(x) \tag{4.5}$$

where f is some real function of the x^α. Let n be the unit vector normal to S, which implies that $n^a = cg^{ab}(dS)_b$ for some factor c. This factor can be determined by the normality condition; assuming that the surface is spacelike, $-1 = n^a n_a = c^2 g^{ab}(dS)_a(dS)_b$, which gives

$$n^a = \pm \frac{g^{ab} dS_b}{\sqrt{-dS_c dS^c}} \tag{4.6}$$

where the $\pm$ sign must be selected based on some choice of orientation. For a Cauchy slice in the diamond, we define $K_{ab} = h^c{}_a \boldsymbol{\nabla}_c n_b$ with an n pointing to the future. The mean curvature is then given by

$$K = \boldsymbol{\nabla}_a n^a = \pm \boldsymbol{\nabla}_a \left(\frac{g^{ab} dS_b}{\sqrt{-dS_c dS^c}} \right) \tag{4.7}$$

In coordinates, $x^\mu = (t, x^\alpha)$, this reads

$$K = \pm \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} \frac{g^{\mu\nu} \partial_\nu S}{\sqrt{-g^{\rho\sigma} \partial_\rho S \partial_\sigma S}} \right) \tag{4.8}$$

where $\sqrt{-g} := \sqrt{-\det(g)}$. From the definition of S, $\partial_\mu S = (1, -\partial_\alpha f)$.

Now consider a null manifold $\mathcal{N}$ (co-dimension 1 in spacetime[14]) whose null (future-pointing) generators are u_+^a. These generators are geodesic, i.e.,

$$u_+^a \boldsymbol{\nabla}_a u_+^b = 0 \tag{4.9}$$

We shall recall here the Gaussian null coordinates [66], characterizing a neighborhood of $\mathcal{N}$. Let λ^+ denote the affine parameter along these geodesics, $d\lambda^+(u_+) = 1$. Let $\mathcal{C}$ be a spatial manifold (co-dimension 2 in spacetime) embedded in $\mathcal{N}$, orthogonal to k_+, and let x^α be coordinates on it. Extend these coordinates x^i to $\mathcal{N}$ by

[14] In this section the spacetime is assumed to be of dimension greater than or equal to 3.

taking x^i constant along the null generators,

$$\pounds_{u_+}(x^i) = 0 \tag{4.10}$$

and consider that $\lambda^+ = 0$ at $\mathcal{C}$. These define a coordinate system $(\lambda^+; x^i)$ on $\mathcal{N}$, and we denote the vector fields tangent to x^i by ∂_i. At every point of $\mathcal{N}$ define the null vector u_- orthogonal to ∂_i, past-pointing, and satisfying the normalization condition

$$(u_-)^a (u_+)_a = 1 \tag{4.11}$$

Extend u_- away from $\mathcal{N}$ by requiring that it is geodesic,

$$u_-^a \boldsymbol{\nabla}_a u_-^b = 0 \tag{4.12}$$

Let λ^- be the corresponding affine coordinate, $d\lambda^-(u_-) = 1$, assumed to vanish at $\mathcal{N}$. Then extend the coordinates $(\lambda^+; x^i)$ away from $\mathcal{N}$ by taking them constant along u_-,

$$\pounds_{u_-}(\lambda^+) = \pounds_{u_-}(x^i) = 0 \tag{4.13}$$

These define coordinates $(\lambda_-, \lambda^+; x^i)$ in a spacetime neighborhood $\mathcal{V}$ of $\mathcal{N}$.

Let us investigate some properties of this coordinate system. First note that, at $\mathcal{N}$, u_+ is everywhere orthogonal to ∂_i. This follows from the fact that $u_+ \cdot \partial_i :=$

$(u_+)_a(\partial_i)^a = 0$ at $\mathcal{C}$ and

$$u_+^a \boldsymbol{\nabla}_a \left(u_{+b}\partial_i^b\right) = u_{+b}u_+^a \boldsymbol{\nabla}_a \partial_i^b = u_{+b}\partial_i^a \boldsymbol{\nabla}_a u_+^b = 0 \tag{4.14}$$

where it was used that u_+ satisfies the geodesic equation, the coordinate condition $\pounds_{u_+}\partial_i = 0$ and that u_+ is everywhere null in $\mathcal{N}$. Second note that since u_- is defined away from $\mathcal{N}$ by the geodesic condition, it is null everywhere in $\mathcal{V}$. Moreover, the inner product between u_- and any other basis vector is constant within $\mathcal{V}$, as follows

$$\begin{aligned} u_-^a \boldsymbol{\nabla}_a \left(u_{-b}u_+^b\right) &= u_{-b}u_-^a \boldsymbol{\nabla}_a u_+^b = u_{-b}u_+^a \boldsymbol{\nabla}_a u_-^b = 0 \\ u_-^a \boldsymbol{\nabla}_a \left(u_{-b}\partial_i^b\right) &= u_{-b}u_-^a \boldsymbol{\nabla}_a \partial_i^b = u_{-b}\partial_i^a \boldsymbol{\nabla}_a u_-^b = 0 \end{aligned} \tag{4.15}$$

where, in each line, we used (in order) the geodesic equation for u_-, the coordinate conditions, $\pounds_{u_-}u_+ = 0$ and $\pounds_{u_-}\partial_i = 0$, and the fact that $u_- \cdot u_- = 0$ in $\mathcal{V}$. Thus, $u_+ \cdot u_- = 1$ and $\partial_i \cdot u_- = 0$ in $\mathcal{V}$. Finally, note that the derivative of $u_+ \cdot u_+$ along u_- vanishes at $\mathcal{N}$,

$$\frac{1}{2}u_-^a \boldsymbol{\nabla}_a(u_{+b}u_+^b) = u_{+b}u_-^a \boldsymbol{\nabla}_a u_+^b = u_{+b}u_+^a \boldsymbol{\nabla}_a u_-^b = u_+^a \boldsymbol{\nabla}_a(u_{+b}u_-^b) - u_{-b}u_+^a \boldsymbol{\nabla}_a u_+^b = 0 \tag{4.16}$$

The reason why this generally only vanishes at $\mathcal{N}$ is because u_+ is typically only geodesic ($u_+^a \boldsymbol{\nabla}_a u_+^b = 0$) at $\mathcal{N}$.

The metric components in this coordinate system thus satisfy the following properties in $\mathcal{V}$,

$$
\begin{aligned}
g_{+-} &= g(u_+, u_-) = 1 \\
g_{--} &= g(u_-, u_-) = 0 \\
g_{-i} &= g(u_-, \partial_i) = 0
\end{aligned}
\tag{4.17}
$$

and the additional properties at $\mathcal{N}$,

$$
\begin{aligned}
g_{++} &= g(u_+, u_+) = 0 \\
g_{+i} &= g(u_+, \partial_i) = 0 \\
\partial_- g_{++} &= u_-^a \boldsymbol{\nabla}_a (u_+ \cdot u_+) = 0
\end{aligned}
\tag{4.18}
$$

We define $h_{ij} := g(\partial_i, \partial_j)$. Since we wish to study the properties of spacelike surfaces approaching $\mathcal{N}$, and the exterior curvature (and the mean curvature) contain one derivative away from the surface, we will consider a first order expansion of the metric in λ^-. Thus, in matrix form, up to first order in λ^-,

$$
g \approx \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 1 & 0 & g_{+i} \\ \hline 0 & g_{+i} & h_{ij} \end{array}\right)
\tag{4.19}
$$

where the components are ordered as $(- + i)$. Note that g_{+i} is first order in λ^-, h_{ij} is zeroth order and g_{++} does not appear since it is second order (due to the last equation in (4.18)). The inverse metric matrix, to first order in λ^-, reads

$$g^{-1} \approx \left(\begin{array}{cc|c} 0 & 1 & -g_{+j}h^{ji} \\ 1 & 0 & 0 \\ \hline -h^{ij}g_{+j} & 0 & h^{ij} \end{array}\right) \tag{4.20}$$

where h^{ij} denotes the inverse of h_{ij}. To first order in λ^-, the determinant is given simply by

$$\det(g) \approx -\det(h) \tag{4.21}$$

where it was used that $\det(g + \delta g) \approx \det(g) + \det(g)\mathrm{tr}(g^{-1}\delta g)$.

Now let us consider a surface $\mathcal{S}$ described by the zero-level of S,

$$S := \lambda^- - \epsilon f(\lambda^+; x^i) \tag{4.22}$$

where ϵ is a (positive) "small" parameter to make it explicit that $\mathcal{S}$ is near $\mathcal{N}$. Suppose that at $\mathcal{C}$, intended to represent a piece of the diamond corner, we have $f(0; x^i) = 0$, indicating that the surface emanates from the corner (as it is intended to represent a portion of a Cauchy slice). In the coordinates constructed above,

$$\partial_\mu S = (1, -\epsilon\partial_+ f, -\epsilon\partial_i f) \tag{4.23}$$

If f is "order 1", then we will be interested in a neighborhood with $\lambda^- \lesssim \epsilon$, so we can use the first-order approximations above to write

$$g^{\mu\nu}\partial_\nu S \approx (-\epsilon\partial_+ f, 1; -h^{ij}g_{+j} - \epsilon h^{ij}\partial_j f) \tag{4.24}$$

where terms such as $\epsilon g_{+i}h^{ij}\partial_j f$, that would appear in the "−" component, are neglected for being of order ϵ^2. Also, we have

$$g^{\mu\nu}\partial_\mu S\partial_\nu S \approx -2\epsilon\partial_+ f \tag{4.25}$$

The mean curvature of $\mathcal{S}$ is therefore

$$K \approx \frac{1}{\sqrt{h}}\partial_\mu\left(\sqrt{h}\frac{g^{\mu\nu}\partial_\nu S}{\sqrt{2\epsilon\partial_+ f}}\right) \tag{4.26}$$

where the + sign was chosen so that $n^\mu = \pm g^{\mu\nu}\partial_\nu S/\sqrt{2\epsilon\partial_+ f}$ points to the future — the reasoning is that, as λ^- grows towards the past (i.e., the interior of the diamond), we need

$$d\lambda^-(n) = n^- = \pm\frac{(-\epsilon\partial_+ f)}{\sqrt{2\epsilon\partial_+ f}} < 0 \tag{4.27}$$

which implies that

$$\pm(-\epsilon\partial_+ f) < 0 \tag{4.28}$$

but in order for $\mathcal{S}$ to be spacelike, $\partial_\mu S\partial^\mu S \approx -2\epsilon\partial_+ f < 0$, which is consistent with the + choice above. Since the argument of the derivative in (4.26) is independent of λ^- in this approximation, we have

$$K \approx \left(\frac{1}{\sqrt{h}}\partial_+\sqrt{h}\right)\frac{1}{\sqrt{2\epsilon\partial_+ f}} + \partial_+\left(\frac{1}{\sqrt{2\epsilon\partial_+ f}}\right) - \frac{1}{\sqrt{h}}\partial_i\left(\sqrt{h}\frac{h^{ij}g_{+j} + \epsilon h^{ij}\partial_j f}{\sqrt{2\epsilon\partial_+ f}}\right) \tag{4.29}$$

Note that the first two terms are order $\epsilon^{-1/2}$, while the third term is order $\epsilon^{1/2}$; therefore the first two terms dominate in the limit $\epsilon \to 0$. In addition, the quantity inside parenthesis in the first term can be identified, in this approximation, with the expansion parameter Θ of the null generators of $\mathcal{N}$, so we have

$$K \approx \frac{\Theta}{\sqrt{2\epsilon\partial_+ f}} + \partial_+\left(\frac{1}{\sqrt{2\epsilon\partial_+ f}}\right) \tag{4.30}$$

Now note that Θ is bounded from above, as follows. Since the corner is smooth and compact, Θ at $\mathcal{C}$ has compact image. Any causal diamond has a compact null horizon, meaning that Θ can evolve with respect to the Raychaudhuri equation. In the present case one can show that Θ must decrease along u^+, so it will either run to $-\infty$ (if a conjugate point appears, i.e., nearby null generators of $\mathcal{N}$ converge to a point), or it may simply stops at a finite value if $\mathcal{N}$ ends before a conjugate point appears (say, when the null generators of $\mathcal{N}$ intersect with another $\mathcal{N}'$ emanating from another portion of the corner). The conclusion is that there exists a finite Θ_0 such that

$$\Theta < \Theta_0 \tag{4.31}$$

and consequently, within the approximations,

$$K < \frac{\Theta_0}{\sqrt{2\epsilon \partial_+ f}} + \partial_+ \left(\frac{1}{\sqrt{2\epsilon \partial_+ f}} \right) \tag{4.32}$$

Now let f be solution of the equation

$$\partial_+ \left(\frac{1}{\sqrt{\partial_+ f}} \right) + \frac{\Theta_0}{\sqrt{\partial_+ f}} = \kappa_0 \tag{4.33}$$

where κ_0 is some (negative) constant. This will imply that

$$K < \frac{\kappa_0}{\sqrt{2\epsilon}} \tag{4.34}$$

so by taking $\epsilon \to 0$ we have that $\mathcal{S}$ will have a mean curvature whose supremum is less than an arbitrarily negative number.

We need to make sure that equation (4.33) has sensible solutions, i.e., representing a spacelike surface within the diamond, for all $\Theta_0 \in \mathbb{R}$ and at least one $\kappa_0 < 0$. The equation is linear in $\zeta(\lambda^+; x) := 1/\sqrt{\partial_+ f}$, and the general solution is

$$\zeta(\lambda^+; x) = \zeta(0; x) e^{-\Theta_0 \lambda^+} + \kappa_0 \int_0^{\lambda^+} d\tau \, e^{-(\lambda^+ - \tau)\Theta_0} \tag{4.35}$$

Then,

$$f(\lambda^+; x) = \int_0^{\lambda^+} d\tau \, \frac{1}{\zeta(\tau; x)^2} \tag{4.36}$$

where it was used that $f(0;x) = 0$. Note that $f(\lambda^+;x) > 0$, which is consistent with the surface being inside the diamond. In order for it to be spacelike we need $\partial_+ f > 0$, which is equivalent to say that $\zeta(\lambda^+, x) > 0$ for λ^+ within its (finite) range. But since $\zeta(0;x) > 0$ can be chosen arbitrarily large (i.e., the surface can be made to start arbitrarily close to being tangent to $\mathcal{N}$), then the term involving κ_0 will not have the opportunity to make ζ become negative. To make this more precise, say that the maximum value of λ^+ at given $x \in \mathcal{C}$ is Λ^+. (In what follows we will omit the x argument.) Now consider three cases, where Θ_0 at x is zero, positive or negative. If $\Theta_0 = 0$, then $\zeta(\lambda^+) = \zeta(0) + \kappa_0\lambda^+$; thus we need $\zeta(0) > -\kappa_0\Lambda^+$. If $\Theta_0 \neq 0$, then $\zeta(\lambda^+) = (\zeta(0) - \kappa_0/\Theta_0)e^{-\Theta_0\lambda^+} + \kappa_0/\Theta_0$; thus if $\Theta_0 < 0$ we need $\zeta(0) > \kappa_0/\Theta_0$, and if $\Theta_0 > 0$, we need $\zeta(0) > -\frac{\kappa_0}{\Theta_0}(e^{\Theta_0\Lambda^+} - 1)$. These impose upper bounds on $\partial_+ f$ at $\mathcal{C}$; moreover, since $\mathcal{C}$ is compact, this bound can be satisfied with $\partial_+ f(0;x)$ strictly positive for all $x \in \mathcal{C}$.

Lastly we must address the fact that the future horizon of the diamond is not a manifold, since it has singularities at the top. In three spacetime dimensions, the singular subset seems to have a (1-dimensional) graph-like shape.[15] Suppose that $\mathcal{C}$ and $\mathcal{C}'$ are disjoint intervals of the corner, and suppose that the null surfaces emanating from them, $\mathcal{N}$ and $\mathcal{N}'$, meet at a line segment $\mathcal{J}$. The surfaces $\mathcal{S}$ and $\mathcal{S}'$, respectively approaching $\mathcal{N}$ and $\mathcal{N}'$ with arbitrarily negative mean curvature, would meet in a singular fashion slightly to the past of $\mathcal{J}$. The idea is to "round off"

[15]The symmetric diamond is exceptional as its future horizon is a cone, with a unique singular point at the top.

this intersection, by interpolating them with a surface $\mathcal{H}$ that is approximately a quadratic surface—a piece of an ellipsoid. We wish to show that if at least one of the radii of the ellipsoid tends to zero, the mean curvature diverges; and, in particular, if it is curved so that the "center" is to the past of the surface, as $\mathcal{H}$ would be, then it diverges negatively. Thus, if we do the rounding off close enough to $\mathcal{J}$, then $\mathcal{S} \cup \mathcal{H} \cup \mathcal{S}'$ will have arbitrarily negative mean curvature.

To define $\mathcal{H}$, consider a coordinate system $(t; x)$ in a neighborhood of $\mathcal{J}$ that is small enough so that the metric can be approximated by the Minkowski metric, $ds^2 = -dt^2 + d(x^1)^2 + d(x^2)^2$.[16] Let $\mathcal{H}$ be described as the zero-level of the function

$$H(t; x) = t - h(x) \tag{4.37}$$

where h is some real function of x. Then formula (4.8) applies, yielding

$$K = \sum_i \partial_i \left(\frac{\partial_i h}{\sqrt{1 - (\partial h)^2}} \right) = \frac{\partial^2 h}{\sqrt{1 - (\partial h)^2}} + \sum_{ij} \frac{\partial_i h \partial_i \partial_j h \partial_j h}{\sqrt{1 - (\partial h)^2}^3} \tag{4.38}$$

where $i, j \in \{1, 2\}$, $(\partial h)^2 := \sum_i \partial_i h \partial_i h$, and $\partial^2 h := \sum_i \partial_i \partial_i h$. The plus sign was chosen so that the normal vector n is future-pointing, $dt(n) > 0$. Now suppose that $\mathcal{H}$ can be approximated by a quadratic surface, meaning

$$h = -\frac{1}{2} \sum_i a_i (x_i)^2 \tag{4.39}$$

[16] The spacetime curvature corrections to this metric will not influence the argument.

for coefficients a_i. We get,

$$K = -\frac{\sum_i a_i}{\sqrt{1-\sum_i a_i^2(x^i)^2}} - \frac{\sum_i (a_i)^3(x^i)^2}{\sqrt{1-\sum_i a_i^2(x^i)^2}^3} \tag{4.40}$$

Note that for this surface to be spacelike we need $(\partial h)^2 = \sum_i a_i^2(x^i)^2 < 1$, which implies a upper bound on the range of x^i. This could be satisfied if we choose $|x^i| < 1/(\sqrt{2}|a_i|)$. Now let us say that x^1 is oriented along $\mathcal{J}$ and x^2 is orthogonal. If $\mathcal{J}$ is a smooth 1-dimensional manifold, and $\mathcal{H}$ is going along it, then a_1 will be bounded while a_2 can be taken to be arbitrarily positive. Therefore the main contribution from the numerators are for $i = 2$, which gives,

$$K \approx -a_2 \frac{1-a_1^2(x^1)^2}{\sqrt{1-\sum_i a_i^2(x^i)^2}^3} < -a_2 \tag{4.41}$$

implying that the mean curvature of $\mathcal{H}$ can be made arbitrarily negative, and consequently the mean curvature of $\mathcal{S} \cup \mathcal{H} \cup \mathcal{S}'$ can be made arbitrarily negative by pushing it towards $\mathcal{N} \cup \mathcal{J} \cup \mathcal{N}'$ with a sharp rounding off at the top.

This argument can be extended to the case where the null surfaces meet at a vertex of the singularity graph, or for the symmetric diamond where the singularity is a point. An analogous argument also applies for the past horizon, showing that a surface with arbitrarily large mean curvature can be constructed.

4.2 The CMC gauge is attainable

We have seen that for any causal diamond in our class of spacetimes (i.e., a state in our system), the CMC condition defines a unique foliation that refers only to intrinsic geometrical structures of the spacetime. This is, therefore, a good non-perturbative prescription to fix the "gauge of time". But why do we have a "gauge of time" in the first place? If we are inclined to regard the diamond as a "self-contained" system, the underlying intuition is that classically the only ingredients we have to construct a diamond spacetime solution is an abstract topological disc and a set of data (h_{ab}, π^{ab}) on it, and there is nothing physically unique about what time coordinate is used to evolve the equations of motion away from the disc, $(h_{ab}, \pi^{ab}) \mapsto (h_{ab}(t), \pi^{ab}(t))$, so one could take $(h_{ab}(t), \pi^{ab}(t))$ as the "initial data" for any t and the resulting solution should correspond to the same physical spacetime. This is essentially the statement that refoliations of the spacetime are gauge. However, from the perspective of the Hamiltonian formalism, we cannot be this vague about what we choose to regard as gauge or not, as they have a sharp definition in terms of the null directions of the symplectic form. If one is determined to ascribe gauge status to certain transformations, it is mandatory to make the appropriate modifications to the theory or specific choices of boundary conditions. As we have discussed in the introduction, we wish to understand the implementation of the CMC gauge fixing program, so we must determine what boundary conditions will imply that refoliations of spacetime are in fact gauge, so that the CMC foliation can be attained via a *gauge transformation*. We will show in this section that this condition is precisely the fixing of the induced boundary metric, $h|_\partial = \gamma$.

As the constraints of General Relativity are of first class, they generate gauge transformations. The Momentum constraint is know to generate spatial diffeomorphisms, while the Hamiltonian constraint generate evolution between Cauchy slices. We are therefore interested in a smearing of the Hamiltonian constraint,

$$H_0[N] := \int d^2x N\sqrt{h}(K^{ab}K_{ab} - K^2 - R + 2\Lambda) \tag{4.42}$$

where the lapse function N vanishes at the boundary, $N|_\partial = 0$, since two Cauchy slices in a causal diamond always meet at the corner. Recall that $\pi^{ab} = \sqrt{h}(K^{ab} - Kh^{ab})$. However, this charge will not generate a gauge transformation if it doesn't generate a (regular) symplectic flow in the first place. Recall that a charge Q generates a symplectic flow X according to the equation $\delta Q = -i_X\Omega$. Now consider the (pre)symplectic form $\Omega = \int_\Sigma d^2x\, \delta\pi^{ab} \wedge \delta h_{ab}$, which is purely a bulk integral. For any (smooth) vector field X in phase space, define $(X^h)_{ab} := i_X\delta h_{ab}$ and $(X^\pi)^{ab} := i_X\delta\pi^{ab}$. We have

$$i_X\Omega = \int_\Sigma d^2x \left[(X^\pi)^{ab}\delta h_{ab} - (X^h)_{ab}\delta\pi^{ab}\right] \tag{4.43}$$

Thus δQ will only be associated with a regular vector field X, via $\delta Q = -i_X\Omega$, if

$$\delta Q = \int_\Sigma d^2x \left[A^{ab}\delta h_{ab} + B_{ab}\delta\pi^{ab}\right] \tag{4.44}$$

for smooth functions A^{ab} and B_{ab}; in particular, δQ must contain no boundary terms involving variations of the dynamical fields. If that is the case, we say that Q is *symplectically differentiable*. The conclusion is that $H_0[N]$ is a gauge-generator if and only if it is symplectically differentiable.

The variation H_0 gives, on-shell,

$$\delta \int d^2x\sqrt{h}\, N(K^{ab}K_{ab} - K^2 - R + 2\Lambda) = \int d^2x\sqrt{h}\, N\delta(K^{ab}K_{ab} - K^2 - R) \tag{4.45}$$

Note that the variation of N itself is irrelevant here since it would end up multiplying a constraint. The variation of $K^{ab}K_{ab} - K^2$ will naturally have the form of a integral of differentials of the dynamical fields over the bulk, so it will not cause issues with differentiability. The only term that could cause issues is

$$\delta R = -\delta h_{ab}R^{ab} + \nabla^a(\nabla^b\delta h_{ab} - h^{bc}\nabla_a\delta h_{bc}) \tag{4.46}$$

since the presence of differentials inside derivatives means that, upon integration by parts and application of Stokes theorem, these differentials may end up in a boundary piece. In fact, the boundary piece one gets is

$$\int d^2x\sqrt{h}\,N(-\delta R) \;\overset{\partial}{\sim}\; -\int_\partial ds\, n^a N(\nabla^b \delta h_{ab} - h^{bc}\nabla_a \delta h_{bc}) \\ + \int_\partial ds\,(n^b \nabla^a N \delta h_{ab} - n_a \nabla^a N h^{bc}\delta h_{bc}) \qquad (4.47)$$

where $\overset{\partial}{\sim}$ means that the two sides differ only by bulk integrals not containing spatial derivatives of field differentials, ds is the proper length measure on the boundary and n^a is the unit normal (outward-pointing) vector to the boundary (and tangent to the Cauchy slice).

Imposing that the lapse function vanishes at the corner, the boundary integral in the first line goes away, leaving

$$\delta H_0 \;\overset{\partial}{\sim}\; \int_\partial ds\left(n^b \nabla^a N \delta h_{ab} - n_a \nabla^a N h^{bc}\delta h_{bc}\right) \qquad (4.48)$$

Let t^a be the unit vector tangent to the corner, so $h_{ab} = n_a n_b + t_a t_b$. Since $N = 0$ along the corner, $t^a \nabla_a N = 0$ which implies that $\nabla^a N = \lambda n^a$ for some scalar λ on the boundary. We therefore get

$$\delta H_0 \;\overset{\partial}{\sim}\; -\int_\partial ds\,\lambda(h^{ab} - n^a n^b)\delta h_{ab} = -\int_\partial ds\,\lambda t^a t^b \delta h_{ab} \qquad (4.49)$$

We see that $H_0[N]$ is symplectically differentiable, for arbitrary lapse satisfying

$N|_\partial = 0$ (so, in particular, arbitrary $\lambda := n^a\nabla_a N \neq 0$), provided that

$$t^a t^b \delta h_{ab} = 0 \tag{4.50}$$

which is equivalent to say that the induced boundary metric on $\partial\Sigma$ must be fixed.

Note that if we do not fix this boundary condition, one can still evolve from one Cauchy slice to another, but the charge generating that trasformation would be an "augmented" version of H_0, where we add to it a boundary term, $Q_\partial := 2\int_\partial ds n^a \nabla_a N$. The boundary term ensures that $H_0 + Q_\partial$ is differentiable (assuming $\delta(n^a\nabla_a N) = 0$) and thus generate a regular symplectic flow. However, on-shell, $H_0 + Q_\partial \approx Q_\partial$, which is not a constant in phase space and thus does not generate a gauge-transformation. In other words, if the Dirichlet boundary condition is not imposed, deformations of the Cauchy slice that tilt the angle that it makes with the corner are non-trivial transformations between distinct physical states. In higher dimensions, the charges generating these corner boosts are given by analogous expressions, but with ds replaced by the induced volume form on the boundary.

4.3 Existence and uniqueness for the Lichnerowicz equation

Here we justify the claim that the Lichnerowicz equation, first appearing in (3.26), always has one and only one solution for any given boundary condition. The proof is a straightforward modification of O'Murchadha and York paper [67] — they consider 3+1 spacetime dimensions, in which case the Lichnerowicz equation is polynomial in the Weyl factor (in fact, this polynomial form is true for all dimensions greater than 2+1). We wish to adapt their argument to 2+1 spacetime dimensions, in which case the Lichnerowicz equation has an exponential form in terms of the Weyl factor. See also [9].

On the spatial disc D, the Lichnerowicz equation associated with pre-initial data $\mathscr{S} = (h_{ab}, \sigma^{ab}, \tau)$ is

$$\nabla^2_{(h)}\phi = R_{(h)} - \sigma^{ab}\sigma_{ab}e^{-\phi} + \chi e^{\phi} \tag{4.51}$$

where $R_{(h)}$ and $\nabla_{(h)}$ are respectively the Ricci scalar and the covariant derivative for the metric h_{ab}. Before proceeding, it is worth noticing that if the Lichnerowicz equation associated with the pre-inital data $\mathscr{S}$ and boundary condition $\phi|_{\partial D} = \varphi$ has a unique solution ϕ_s, then the Lichnerowicz equation for a Weyl-tranformed data $\mathscr{S}' = (e^{\lambda}h_{ab}, e^{-2\lambda}\sigma^{ab}, \tau)$ and boundary condition $\phi'|_{\partial D} = \varphi - \lambda$ also has a unique solution given by $\phi'_s = \phi_s - \lambda$. This can be easily seen from the transformation properties of the Laplacian and the Ricci scalar, as discussed around equation (3.28), since the Lichnerowicz equation for $\mathscr{S}'$ can be re-expressed as

$$\nabla^2_{(h)}\phi' = \left(R_{(h)} - \nabla^2_{(h)}\lambda\right) - e^{-\lambda}\sigma^{ab}\sigma_{ab}e^{-\phi'} + e^{\lambda}\chi e^{\phi'} \tag{4.52}$$

revealing that $\phi' + \lambda$ satisfies the same equation as ϕ, with the same boundary condition.

To prove uniqueness of solution, suppose that (4.51) has a solution ϕ_s and consider the transformation in which $\lambda = \phi_s$, so that the Lichnewowicz equation for ϕ', as in (4.52), becomes

$$\nabla^2_{(h)}\phi' = e^{-\lambda}\sigma^{ab}\sigma_{ab}\left(1 - e^{-\phi'}\right) + e^{\lambda}\chi\left(e^{\phi'} - 1\right) \tag{4.53}$$

with vanishing boundary condition for ϕ'. As expected, $\phi' = 0$ is a solution. Now multiply this equation by ϕ' on both sides and integrate over Σ, with respect to the volume form associated with h_{ab},

$$\int \phi'\nabla^2_{(h)}\phi' = \int \left[e^{-\lambda}\sigma^{ab}\sigma_{ab}\phi'\left(1 - e^{-\phi'}\right) + e^{\lambda}\chi\phi'\left(e^{\phi'} - 1\right)\right] \geq 0 \tag{4.54}$$

observing that the right-hand side is nonnegative since $\phi'(1 - e^{-\phi'})$ and $\phi'(e^{\phi'} - 1)$ are nonnegative for any ϕ', and $\sigma^{ab}\sigma_{ab}$ and χ are also nonnegative. But integrating the left-hand side by parts gives

$$\int \phi'\nabla^2_{(h)}\phi' = \int \nabla_{(h)} \cdot \left(\phi'\nabla_{(h)}\phi'\right) - \int \left(\nabla_{(h)}\phi'\right)^2 \leq 0 \tag{4.55}$$

where the first term vanishes by using Stokes' theorem and imposing $\phi' = 0$ at the boundary. We therefore conclude that $\nabla_{(h)}\phi' = 0$ and, given the boundary condition, also that $\phi' = 0$. That is, this equation has no solutions other than $\phi' = 0$ and thus (4.51) has no solutions other than ϕ_s.

In O'Murchadha and York paper, two proofs of existence are offered. Here we shall only revisit one of the proofs, based on the construction of a sequence of functions that converge to the solution. To unclutter the notation, let us write (4.51) in its general form

$$\nabla^2_{(h)}\phi = F(\phi; x)\,, \quad \phi|_{\partial D} = \varphi \tag{4.56}$$

where

$$F(\phi; x) = c(x) - a(x)e^{-\phi} + b(x)e^{\phi} \tag{4.57}$$

in which $a(x)$ and $b(x)$ are nonnegative functions on Σ, and $c(x)$ is a function on Σ. As explained, for the purposes of proving existence and uniqueness of solutions, we can always perform a transformation $c \to c - \nabla^2\lambda$, $a \to ae^{-\lambda}$, $b \to be^{\lambda}$ and $\varphi \to \varphi - \lambda|_{\partial D}$ (which is the form of the problem expressed in (4.52)). With this freedom, φ and c can be chosen to be any given functions, according to convenience, and a and b will fall into one of the three cases below:

(i) $a \geq 0$ and $b > 0$. This is the case of main interest, when $\sigma^{ab}\sigma_{ab} \geq 0$ and $\chi > 0$. For reasons that will become clear, it is convenient to choose c to be any negative constant, $c_0 < 0$.

(ii) $a \geq 0$ and $b = 0$. This is the case for the maximal slice ($\tau = 0$) with zero cosmological constant. Here we can choose c as any function such that $c(x) > 0$ if $a(x) > 0$ and $c(x) = 0$ if $a(x) = 0$.

(iii) $a = 0$ and $b = 0$. The existence of solution for this case is trivial since we

can choose $c = 0$ so the equation reduces to $\nabla^2\phi = 0$.

In all cases we could choose, e.g., $\varphi = 0$, but this would not amount to any real simplification.

The core of the proof lies in the existence of functions ϕ_+ and ϕ_- on Σ, with $\phi_+ \geq \phi_-$, such that

$$F(\phi_+) \geq 0\,,\ \ \phi_+|_{\partial D} \geq \varphi$$

$$F(\phi_-) \leq 0\,,\ \ \phi_-|_{\partial D} \leq \varphi$$

and we can see that this is true for cases (i) and (ii) above. For case (i), the b term dominates as long as ϕ_+ is sufficiently large, making $F > 0$; and for ϕ_- very negative the b term is suppressed and the a term dominates wherever $a > 0$, while the $c := c_0 < 0$ term ensures that $F < 0$ even in regions where $a = 0$. For case (ii), the $c(x)$ term dominates when ϕ_+ is very large, making $F \geq 0$; while the a term dominates when ϕ_- is very negative, making $F \leq 0$. It will be convenient (and it is possible) to take ϕ_- and ϕ_+ to be constants satisfying $\phi_-|_{\partial D} \leq \varphi \leq \phi_+|_{\partial D}$.

Note that if we had a "case (iv)" in which $a \geq 0$ and $b < 0$ then it would not be generally possible to find ϕ_+ and ϕ_- satisfying the desired conditions. This is would be the case for a positive cosmological constant, as $\chi = -2\Lambda + \tau^2/2 < 0$ for τ in some interval around 0, i.e., around the maximal slice. The problem occurs at any points where σ_{ab} vanishes, so that $a = 0$. At these points, $F = c + be^{\phi}$ is a monotonically decreasing function of ϕ, so it is not possible to have both $\phi_+ \geq \phi_-$

and $F(\phi_-) \leq 0 \leq F(\phi_+)$.

The proof begins by constructing a sequence of functions ϕ_n recurssively defined by

$$\nabla^2 \phi_n - \kappa \phi_n = F(\phi_{n-1}) - \kappa \phi_{n-1} \tag{4.58}$$

with boundary condition $\phi_n|_{\partial D} = \varphi$, where

$$\kappa := \max_{x \in D} \max_{\phi_- \leq \phi \leq \phi_+} \left| \frac{\partial F}{\partial \phi}(\phi; x) \right| \tag{4.59}$$

and the starting function is $\phi_0 := \phi_+$. Note that κ is well defined because, for each x, $\phi(x)$ is in the compact interval $[\phi_-(x), \phi_+(x)]$, and then x is maximized over the compact set D. We will omit the reference to the metric in $\nabla = \nabla_{(h)}$ because it only matters to us that the metric is Riemannian, as this ensures that the differential operator in (4.58), $L := \nabla^2 - \kappa$, is strictly elliptic[17]. The goal is to prove that this sequence is monotonically decreasing, in the sense that $\phi_n < \phi_{n-1}$, and also bounded from below by ϕ_-. This ensures that the sequence converges and, by a simple argument, that the limit is the solution of (4.56).

First let us see that, if the limit exists, it should be the solution of (4.56). As a note of consistency, observe that the solution is a fixed point of the iteration. That is, if ϕ_{n-1} is a solution for some $n > 1$, $\nabla^2 \phi_{n-1} = F(\phi_{n-1})$, then $\nabla^2(\phi_n - \phi_{n-1}) - \kappa(\phi_n - \phi_{n-1}) = 0$, with boundary condition $(\phi_n - \phi_{n-1})|_{\partial D} = \varphi - \varphi = 0$, and it is

[17] A linear second order differential operator $L = A_{ij}(x)\partial_i \partial_j + B_i(x)\partial_i + C(x)$ is called *strictly elliptic* on a domain Ω if there is $\lambda > 0$ such that $A_{ij}(x)\xi_i\xi_j \geq \lambda \xi_i \xi_i$, $x \in \Omega$, for all vectors ξ. (Sum over repeated indices implied.)

clear that $\phi_n = \phi_{n-1}$. Now we proceed with the proper proof. Suppose that ϕ is the limit of the sequence $\{\phi_n\}$, meaning that for every $\epsilon > 0$ there exists N such that $|\phi - \phi_n| < \epsilon$ for all $n > N$. This assumes that the convergence is uniform, which is guaranteed by Dini's theorem as long as the limit ϕ is continuous[18]. For $n > N + 1$ we have $|\phi_{n-1} - \phi_n| < 2\epsilon$, so subtracting $F(\phi_n)$ on both sides of (4.58) we get

$$\begin{aligned} |\nabla^2\phi_n - F(\phi_n)| &= |\kappa(\phi_n - \phi_{n-1}) + F(\phi_{n-1}) - F(\phi_n)| \\ &< \kappa|\phi_n - \phi_{n-1}| + |F(\phi_{n-1}) - F(\phi_n)| \\ &< 2\kappa|\phi_n - \phi_{n-1}| \\ &< 4\kappa\epsilon \end{aligned} \tag{4.60}$$

showing that ϕ_n converges (weakly[19]) to a solution of $\nabla^2\phi - F(\phi) = 0$.

Now let us show that the sequence of solutions ϕ_n is monotonically decreasing and bounded from below. We start by reviewing a important result from the theory of linear elliptic differential equations [68]. Given a strictly elliptic differential operator L, as defined in footnote 17, with $C(x) \leq 0$, and given a function f, let u be the solution of $-L[u] = f$ with boundary condition $u|_{\partial D} = \varphi$. A func-

[18] *Dini's theorem* states that if a monotonically increasing (or decreasing) sequence of continuous real functions on a compact topological space converges pointwise to a continuous function, then the convergence is uniform.

[19] In the theory of differential equations, a function u is said to be a weak solution of the differential operator E on a domain D if $\int_D E[u]\,\xi = 0$ for all smooth functions ξ supported on arbitrary compact subsets of $\mathrm{Int}(D)$; it is assumed that integrations by parts have been formally applied to the integral so as to move all the derivatives from u to ξ (thus u may be a weak solution for E even if it does not have well-defined derivatives). For example, a weak solution ϕ for (4.56) is required to satisfy $\int(\nabla^2\phi - F(\phi))\xi = \int(\phi\nabla^2\xi - F(\phi)\xi) = 0$ for all smooth ξ that vanishes in a neighborhood of the boundary of the disc. The result above shows that $\nabla^2\phi_n - F(\phi_n)$ converges to zero, which implies that ϕ_n converges to a weak solution of (4.56).

tion u^+ is called a *supersolution* of this differential equation if $-L[u^+] \geq f$ and $u^+|_{\partial D} \geq \varphi$; similarly, a function u^- is called a *subsolution* of this differential equation if $-L[u^-] \leq f$ and $u^-|_{\partial D} \leq \varphi$. If u^+ is a supersolution then $u \leq u^+$, and if u^- is a subsolution then $u^- \leq u$. The equation (4.58) defining ϕ_n satisfies the conditions of the theorem since $L = \nabla^2 - \kappa$ is strictly elliptic with $C = -\kappa < 0$, and here $f = f_{n-1} := -F(\phi_{n-1}) + \kappa\phi_{n-1}$. The simplest way to construct a supersolution is to consider a constant function ϕ_n^+, in which case $-L[\phi_n^+] = \kappa\phi_n^+$, so we can take

$$\phi_n^+ = \max(\kappa^{-1} f_{n-1}, \varphi) \tag{4.61}$$

where the maximization also runs over D. (Note that $\kappa \neq 0$, except in case (*iii*) above, which is trivial.) Similarly, we can construct a subsolution by taking

$$\phi_n^- = \min(\kappa^{-1} f_{n-1}, \varphi) \tag{4.62}$$

Therefore, the (sub)supersolution theorem implies that

$$\min(\kappa^{-1} f_{n-1}, \varphi) \leq \phi_n \leq \max(\kappa^{-1} f_{n-1}, \varphi) \tag{4.63}$$

As mentioned before, the starting point is $\phi_0 := \phi_+$. In the equation defining ϕ_1 we have $\kappa^{-1} f_0 = -\kappa^{-1} F(\phi_+) + \phi_+ \leq \phi_+$, where we have used that $F(\phi_+) \geq 0$. Thus, since $\varphi \leq \phi_+$, it follows from (4.63) that $\phi_1 \leq \phi_+ =: \phi_0$. Also, as $\phi_- \leq \varphi \leq \phi_+$

and $F(\phi_-) \leq 0 \leq F(\phi_+)$, we have $F(\phi_+) - F(\phi_-) \leq \kappa(\phi_+ - \phi_-)$, so that $\kappa^{-1} f_0 = -\kappa^{-1}F(\phi_+) + \phi_+ \geq -\kappa^{-1}F(\phi_-) + \phi_- \geq \phi_-$ and it follows from (4.63) that $\phi_1 \geq \phi_-$.

Having established that $\phi_- \leq \phi_1 \leq \phi_0 = \phi_+$, we now proceed by induction: suppose that $\phi_- \leq \phi_{n-1} \leq \phi_{n-2} \leq \phi_+$ for $n > 1$ and show that $\phi_- \leq \phi_n \leq \phi_{n-1} \leq \phi_+$. Replicating the argument above, let us first prove that $\phi_n \geq \phi_-$. In the equation defining ϕ_n we have $\kappa^{-1} f_{n-1} = -\kappa^{-1}F(\phi_{n-1}) + \phi_{n-1}$, so using $|F(\phi_{n-1}) - F(\phi_-)| \leq \kappa(\phi_{n-1} - \phi_-)$ we see that $\kappa^{-1} f_{n-1} \geq -\kappa^{-1}F(\phi_-) + \phi_- \geq \phi_-$. Hence $\phi_n \geq \min(\kappa^{-1} f_{n-1}, \varphi) \geq \min(\phi_-, \varphi) = \phi_-$. Next we prove that $\phi_n \leq \phi_{n-1}$. Subtracting equation (4.58) for ϕ_n from that for ϕ_{n-1} we get

$$\nabla^2(\phi_n - \phi_{n-1}) - \kappa(\phi_n - \phi_{n-1}) = F(\phi_{n-1}) - F(\phi_{n-2}) + \kappa(\phi_{n-2} - \phi_{n-1}) \qquad (4.64)$$

This equation satisfies the super/subsolution theorem for the same operator $L = \nabla^2 - \kappa$, where here we have $u = \phi_n - \phi_{n-1}$ and $f = F(\phi_{n-2}) - F(\phi_{n-1}) - \kappa(\phi_{n-2} - \phi_{n-1})$. Since $|F(\phi_{n-2}) - F(\phi_{n-1})| \leq \kappa(\phi_{n-2} - \phi_{n-1})$ we see that $f \leq 0$. Moreover, for $n > 1$, $(\phi_n - \phi_{n-1})|_{\partial D} = \varphi - \varphi = 0$, implying that $\max(\kappa^{-1} f, u|_{\partial D}) = 0$ and thus $\phi_n - \phi_{n-1} \leq 0$. This concludes the proof that $\phi_- \leq \cdots \leq \phi_n \leq \phi_{n-1} \leq \cdots \leq \phi_+$.

5 Quotienting by conformal transformations

In this section we will fill in some of the details left out in Sec. 3.2. In particular we show that the reduced phase space is isomorphic to the cotangent bundle of the space of conformal isometries of the disk (subjected to the appropriate boundary conditions) and that this space of conformal geometries is isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. This identification of the phase space with the cotangent bundle of space of conformal geometries of the Cauchy slice is known in the literature and quite generic in the context of gravity [9, 10], but our presentation is slightly different and specialized to the case of our interest.

5.1 The cotangent bundle of the space of conformal geometries

We have explained in Sec. 3.2 that the reduced phase space for the causal diamonds is, as given in (3.34), isomorphic to the space of equivalence classes

$$\widetilde{\mathcal{P}} = \{(h_{ab}, \sigma^{ab}) \sim (\Psi_* e^{\lambda} h_{ab}, \Psi_* e^{-2\lambda} \sigma^{ab})\} \tag{5.1}$$

where $\Psi : \Sigma \to \Sigma$ is a boundary-trivial diffeomorphism of the disc (i.e., Ψ acts as the identity on the boundary of the disc) and $\lambda : \Sigma \to \mathbb{R}$ vanishes at the boundary of the disc. Here we shall describe the cotangent bundle of the space of conformal geometries of the disk in a way that makes explicit that it is isomorphic to $\widetilde{\mathcal{P}}$.

First let us describe the cotangent bundle of a quotient space in general terms. That is, given a manifold $\mathcal{M}$ and a Lie group G that acts on $\mathcal{M}$, we wish to characterize the cotangent bundle of the space of orbits, $T^*(\mathcal{M}/G)$, in terms of the cotangent bundle of the manifold, $T^*\mathcal{M}$. The goal is to later particularize to the case where $\mathcal{M}$ is the space of Riemmanian metrics on a disc and G is the group of boundary-trivial conformal transformations. We shall assume that $\mathcal{M}/G$ has a manifold structure, which is guaranteed [69, 70] to happen (at least in finite dimensions) when the stabilizer group of this action is a closed subgroup of G. Let us think of $\mathcal{M}$ as a principal G-bundle over $\mathcal{M}/G$, denoting by $p : \mathcal{M} \to \mathcal{M}/G$ the quotient map that takes a point $x \in \mathcal{M}$ to the equivalence class $[x] := \{x \sim \Lambda_g x; g \in G\}$, where Λ_g is the diffeomorphism on $\mathcal{M}$ produced by $g \in G$. Since p is a projection map, the push-forward p_* is a surjection from $T_x\mathcal{M}$ onto $T_{[x]}(\mathcal{M}/G)$. Also, if two vectors ξ and ξ' at x project to the same vector under p_*, then they must differ by an element of the kernel of p_*, i.e., $\xi' = \xi + \eta$ for some η satisfying $p_*\eta = 0$. This means that $T_{[x]}(\mathcal{M}/G)$ is isomorphic to the quotient of $T_x\mathcal{M}$ by $\ker(p_*)$, where two vectors at $x \in \mathcal{M}$ are identified if they differ by an element of $\ker(p_*)$. In display,

$$T_{[x]}(\mathcal{M}/G) = T_x\mathcal{M}/\ker(p_*) \tag{5.2}$$

There is also another way to characterize $T_{[x]}(\mathcal{M}/G)$ by noticing that vectors at any points along the fiber over $[x]$ all project to vectors at the same base point $[x]$. We have $p \circ \Lambda_g = p$, for all $g \in G$, as p projects the whole fiber to the same base point. This implies that $p_*\Lambda_{g*} = p_*$ and therefore the vector ξ at x has the same image under p_* as the vector $\Lambda_{g*}\xi$ at $\Lambda_g x$. Moreover, since Λ_g is a diffeomorphism of $\mathcal{M}$, its derivative Λ_{g*} induces an isomorphism between the tangent spaces $T_x\mathcal{M}$ and $T_{\Lambda_g x}\mathcal{M}$, and this isomorphism preserves the kernel of p_*. Thus, if ξ at x projects to a given vector at $[x]$, then the only other vectors on $T\mathcal{M}$ that project to the same vector are given by $\Lambda_{g*}(\xi + \eta)$ for all $g \in G$ and all $\eta \in \ker(p_*)$. In display, the tangent bundle of $\mathcal{M}/G$ is identified with the following quotient of the tangent bundle of $\mathcal{M}$

$$T(\mathcal{M}/G) = \{\xi \sim \Lambda_{g*}(\xi + \eta);\, g \in G,\, \xi \in T\mathcal{M} \text{ and } \eta \in \ker(p_*)\} \tag{5.3}$$

where the projection on the tangent bundle of $\mathcal{M}/G$, $\widetilde{\pi} : T(\mathcal{M}/G) \to \mathcal{M}/G$, is simply given by $\widetilde{\pi}([\xi]) = p(\pi(\xi))$, where $\pi : T\mathcal{M} \to \mathcal{M}$ is the projection on the tangent bundle of $\mathcal{M}$ and $\xi \in T\mathcal{M}$ is any representative of the class $[\xi] \in T(\mathcal{M}/G)$.

Now, most importantly, let us describe the cotangent bundle of $\mathcal{M}/G$. If $\widetilde{\alpha}$ is a 1-form at $[x] \in \mathcal{M}/G$, then its pull-back to $x \in \mathcal{M}$ via p has the property of annihilating the whole kernel of p_*,

$$p^*\widetilde{\alpha}(\eta) = \widetilde{\alpha}(p_*\eta) = \widetilde{\alpha}(0) = 0 \tag{5.4}$$

where $\eta \in \ker(p_*)$. On the other direction, every 1-form α at $x \in \mathcal{M}$ that annihilates $\ker(p_*)$ defines a 1-form at $[x] \in \mathcal{M}/G$. This can be seen from the characterization of vectors at $[x]$ given in (5.2), since such an α defines a linear map to $\mathbb{R}$ which satisfies $\alpha(\xi) = \alpha(\xi + \eta)$ for every vector ξ at x and $\eta \in \ker(p_*)$, and therefore is a well-defined linear map from $T_{[x]}(\mathcal{M}/G)$ to $\mathbb{R}$. Moreover, if we denote by $\widetilde{\alpha}$ the 1-form at $[x]$ defined in this way from α, it is clear that $\alpha = p^*\widetilde{\alpha}$, showing that there is an isomorphism

$$T^*_{[x]}(\mathcal{M}/G) = \{\alpha \in T^*_x\mathcal{M};\ \text{where } \alpha(\eta) = 0 \text{ for all } \eta \in \ker(p_*)\} \tag{5.5}$$

Similarly to the other characterization of the tangent bundle, 1-forms on $\mathcal{M}/G$ can also be described in terms of a quotient over 1-forms at different points along the fiber $p^{-1}([x])$. From the identity $p \circ \Lambda_{g^{-1}} = p$ we have that $\Lambda^*_{g^{-1}}p^* = p^*$, and it follows that the $\ker(p_*)$-annihilating 1-forms at x are related via $\Lambda^*_{g^{-1}}$ to the $\ker(p_*)$-annihilating 1-forms at $\Lambda_g x$. In fact, if η is any vector in $\ker(p_*)$ at $\Lambda_g x$ then $\Lambda_{g^{-1}*}\eta$ is in $\ker(p_*)$ at x; so if α annihilates $\ker(p_*)$ at x then $0 = \alpha(\Lambda_{g^{-1}*}\eta) = \Lambda^*_{g^{-1}}\alpha(\eta)$, revealing that $\Lambda^*_{g^{-1}}\alpha$ annihilates $\ker(p_*)$ at $\Lambda_g x$. In addition, if α at x is related to $\widetilde{\alpha}$ at $[x]$, then $\Lambda^*_{g^{-1}}\alpha$ are also related to the same $\widetilde{\alpha}$, for all $g \in G$. This gives the characterization of the cotangent bundle of $\mathcal{M}/G$ as the following set of equivalence classes

$$T^*(\mathcal{M}/G) = \{\alpha \sim \Lambda^*_{g^{-1}}\alpha; g \in G,\ \text{where } \alpha \in T^*\mathcal{M} \text{ satisfies } \alpha(\ker(p_*)) = 0\} \tag{5.6}$$

Now we can particularize the results above to the case of interest, where $\mathcal{M}$ is taken to be space $\text{Riem}(D;\gamma)$ of Riemmanian metrics h on the disc D satisfying the desired boundary condition for the induced boundary metric, $h|_\partial = \gamma$, and G is the group of boundary-trivial conformal transformations on the metric. The topology of $\text{Riem}(D;\gamma)$ can be defined by seeing it as an (open) subset of the vector space

$\mathrm{Sym}^0_2(D)$ of symmetric type-$\binom{0}{2}$ tensors on D (which can be assumed to be some Sobolev space, although we shall not worry about these details). Naturally a tangent vector ξ at any point $h \in \mathrm{Riem}(D;\gamma)$ can be identified with $\xi_{ab} \in \mathrm{Sym}^0_2(D^*)$, where $\mathrm{Sym}^0_2(D^*)$ is the subset of $\mathrm{Sym}^0_2(D)$ for which $\xi_{ab}t^a t^b = 0$ for any vector t^a tangent to the boundary S^1. The tangent bundle of $\mathrm{Riem}(D;\gamma)$ is trivial and equal to

$$T[\mathrm{Riem}(D;\gamma)] = \mathrm{Riem}(D;\gamma) \times \mathrm{Sym}^0_2(D^*) \tag{5.7}$$

Dual vectors α at any point $h \in \mathrm{Riem}(D;\gamma)$, taken to be continuous linear functions from vectors to $\mathbb{R}$, can be naturally identified with symmetric type-$\binom{2}{0}$ tensors α_{ab} on D. The pairing will be defined as

$$\alpha(\xi) := \int \vartheta_h \alpha^{ab} \xi_{ab} \tag{5.8}$$

where ϑ_h is the volume form associated with h.[20] The cotangent bundle has the trivial structure

$$T^*[\mathrm{Met}(\mathcal{M})] = \mathrm{Riem}(D;\gamma) \times \mathrm{Sym}^2_0(D) \tag{5.9}$$

The group of boundary-trivial conformal transformations, $G = \mathrm{Con}(D^*)$, act on the metric as $\Lambda_{(\Psi,\Omega)}h := \Psi_*\Omega h$, where Ψ is a diffeomorphism of the disc acting as the identity at the boundary and Ω is a positive real function that equals 1 at the

[20] As the boundary S^1 has measure zero in this integration, the boundary condition on the vectors implies no constraints on the dual vectors, which is why they are identified with matrices in $\mathrm{Sym}^2_0(D)$ instead of $\mathrm{Sym}^2_0(D^*)$.

boundary. The kernel of p_* is composed of vectors tangent to the orbits of this action, i.e., the vectors in $\mathrm{Riem}(D;\gamma)$ induced from the algebra of G. That is, if X is an element of the algebra of G, the induced vector η^X is

$$\eta^X_{ab} := \frac{\partial}{\partial t}\Lambda_{\exp(tX)}h_{ab}\Big|_{t=0} \tag{5.10}$$

and this is in $\ker(p_*)$. We can separately consider vectors induced from infinitesimal diffeomorphisms and from infinitesimal Weyl scalings. Let us begin with (boundary-trivial) Weyl scalings, where a typical algebra element is denoted by W and the exponentiation defines a curve on the group $\exp(tW) = (I, e^{tw})$, in which w is a real function on the disc vanishing at the boundary. The induced vector at a point h in $\mathrm{Riem}(D;\gamma)$ is

$$\eta^W_{ab} = \frac{\partial}{\partial t}e^{tw}h_{ab}\Big|_{t=0} = wh_{ab} \tag{5.11}$$

Next, an infinitesimal (boundary-trivial) diffeomorphism is labeled by a vector field V on the disc, vanishing at the boundary, and the exponential defines a curve on the group $\exp(tV) = (\Psi_{tV}, 1)$, in which Ψ_{tV} is the diffeomorphism defined by running along the integral curves of V for a parameter length t. The induced vector at a point h in $\mathrm{Riem}(D;\gamma)$ is

$$\eta^V_{ab} = \frac{\partial}{\partial t}\Psi_{tV*}h_{ab}\Big|_{t=0} = -\pounds_V h_{ab} = -2\nabla_{(a}V_{b)} \tag{5.12}$$

where ∇ is the derivative associated with h. The set of vectors η^W and η^V, for all W and V, spans the kernel of p_*. With this we can determine the space of $\ker(p_*)$-annihilating 1-forms at h, thus providing a characterization for 1-forms at $[h]$ according to (5.5). So

$$0 = \alpha(\eta^W) = \int \vartheta_h \alpha^{ab} w h_{ab} \tag{5.13}$$

for all w implies that α must be traceless with respect to h, i.e., $\alpha^{ab} h_{ab} = 0$; and

$$0 = \alpha(\eta^V) = -2 \int \vartheta_h \alpha^{ab} \nabla_{(a} V_{b)} = 2 \int \vartheta_h \nabla_a \alpha^{ab} V_b \tag{5.14}$$

for all V implies that α must be divergenceless with respect to h, i.e., $\nabla_a \alpha^{ab} = 0$. Note that there is no boundary term in the integration by parts performed above because V vanishes at the boundary. Those conditions on α are clearly familiar to us: if α^{ab} is interpreted as an extrinsic curvature on the disc, the traceless condition would correspond to the gauge-fixing of time by CMC surfaces (which leaves only the traceless part of the extrinsic curvature as dynamical), and the divergenceless condition would correspond to the momentum constraint.

Finally, we are interested in describing the cotangent bundle of $\mathrm{Riem}(D; \gamma)/\mathrm{Con}(D^*)$ globally according to (5.6). To this end, it only remains to compute the pullback of 1-forms by $\Lambda_{(\Psi,\Omega)}$. But first we should compute the push-forward of a vector ξ_{ab} by $\Lambda_{(\Psi,\Omega)}$. Naturally ξ_{ab} is tangent at h_{ab} to the curve $h_{ab} + t\xi_{ab}$, so its push-forward should be tangent to the curve $\Psi_*\Omega(h_{ab}+t\xi_{ab})$, that is,

$$\Lambda_{(\Psi,\Omega)*}\xi_{ab} = \Psi_*\Omega\xi_{ab} \tag{5.15}$$

If α is a 1-form at h, its pull back $\Lambda^*_{(\Psi,\Omega)}\alpha$ is based at $\Lambda^{-1}_{(\Psi,\Omega)}h = \Omega^{-1}\Psi_*^{-1}h$ and satisfies

$$\Lambda^*_{(\Psi,\Omega)}\alpha(\xi) = \alpha(\Lambda_{(\Psi,\Omega)*}\xi) = \int \vartheta_h \alpha^{ab}\Psi_*\Omega\xi_{ab} \tag{5.16}$$

where here ξ is a vector at $\Lambda^{-1}_{(\Psi,\Omega)}h$. Using that $\vartheta_{\Omega^{-1}\Psi_*^{-1}h} = \Omega^{-1}\Psi_*^{-1}\vartheta_h$, we can rewrite the last integral as

$$\int \vartheta_h \alpha^{ab}\Psi_*\Omega\xi_{ab} = \int \Psi_*\left[\vartheta_{\Omega^{-1}\Psi_*^{-1}h}(\Omega^2\Psi_*^{-1}\alpha^{ab})\xi_{ab}\right] \tag{5.17}$$

and from the diffeomorphism invariance of integrals,

$$\Lambda^*_{(\Psi,\Omega)}\alpha(\xi) = \alpha(\Lambda_{(\Psi,\Omega)*}\xi) = \int \vartheta_{\Omega^{-1}\Psi_*^{-1}h}(\Omega^2\Psi_*^{-1}\alpha^{ab})\xi_{ab} \tag{5.18}$$

which allows us to read off

$$(\Lambda^*_{(\Psi,\Omega)}\alpha)^{ab} = \Omega^2\Psi_*^{-1}\alpha^{ab} \tag{5.19}$$

or, taking the inverse transformation,

$$(\Lambda^*_{(\Psi,\Omega)^{-1}}\alpha)^{ab} = \Psi_*\Omega^{-2}\alpha^{ab} \tag{5.20}$$

Therefore, as a point in $T^*\text{Riem}(D;\gamma)$ is labeled by the pair (h_{ab}, α^{ab}), the characterization (5.6) for the cotangent bundle of $\text{Riem}(D;\gamma)/\text{Con}(D^*)$ gives

$$\begin{aligned} T^*(\text{Riem}(D;\gamma)/\text{Con}(D^*)) \quad = \quad & \{(h_{ab}, \alpha^{ab}) \sim (\Psi_*\Omega h_{ab}, \Psi_*\Omega^{-2}\alpha^{ab}); \\ & (\Psi, \Omega) \in \text{Con}(D^*), \text{ where } (h_{ab}, \alpha^{ab}) \in T^*\text{Riem}(D;\gamma) \\ & \text{satisfies } h_{ab}\alpha^{ab} = 0 \text{ and } \nabla_a^{(h)}\alpha^{ab} = 0\} \end{aligned} \tag{5.21}$$

This is precisely the characterization of the reduced phase space for the causal diamonds displayed in (3.34), which proves the claim that

$$\widetilde{\mathcal{P}} = T^*(\text{Riem}(D;\gamma)/\text{Con}(D^*)) \tag{5.22}$$

We stress that this result relies on that fact that the group of conformal transformations acts trivially at the boundary, as it was important that there was no boundary term coming from Stokes' theorem in (5.14).

5.2 The space of conformal geometries on a disc

Now we wish to provide a more direct characterization for the configuration space, i.e., the space of conformal geometries on a disc $\mathcal{Q} := \text{Riem}(D;\gamma)/\text{Con}(D^*)$. According to the *Riemann mapping theorem*, in complex analysis, given any non-empty simply-connected open subset of the complex plane $\mathbb{C}$, which is not all of $\mathbb{C}$, there exists a biholomorphic map onto the open unit disc $\mathbb{D} := \{z \in \mathbb{C}, |z| < 1\}$. The

Carathéodory theorem extends this theorem to closed sets, stating that if f maps the open unit disc $\mathbb{D}$ conformally onto the (bounded) open set $U \subset \mathbb{C}$, then f has a continuous one-to-one extension to a function from the closure of $\mathbb{D}$, $\overline{\mathbb{D}}$, onto the closure of U, $\overline{U}$, if (and only if) ∂U is a Jordan curve[21]. This can be applied to our problem, yielding that for any two matrics h and h' there exists an orientation-preserving diffeomorphism Ψ and a Weyl factor Ω such that $h' = \Psi_* \Omega h$. As before, we shall denote a general conformal transformation by the pair (Ψ, Ω).

This means that if we had no constraint on the form of (Ψ, Ω) at the boundary, all metrics on the disc would be conformally equivalent. In other words, $\mathrm{Riem}(D;\gamma)/\mathrm{Con}(D)$ consists of a single point. However, because of the constraints, some states will no longer be conformally connected, so $\mathrm{Riem}(D;\gamma)/\mathrm{Con}(D^*)$ will be non-trivial. In order to determine it, note that $\mathcal{Q} = \mathrm{Riem}(D;\gamma)/\mathrm{Con}(D^*)$ is a homogeneous space for the group $\mathit{Diff}^+(S^1)$. Namely, given a diffeomorphism acting on the boundary of the disc $\psi \in \mathit{Diff}^+(S^1)$, define its left action on a conformal geometry $[h] \in \mathrm{Riem}(D;\gamma)/\mathrm{Con}(D^*)$ by

$$\psi[h] := [\Psi_* \Omega h] \tag{5.23}$$

where $\Psi \in \mathrm{Diff}^+(D)$ is any extension of ψ to the interior of the disc, so that $\Psi|_\partial = \psi$

[21]A *Jordan curve* is a simple closed curve, i.e., a non-self-intersecting continuous loop on the plane.

and Ω is any positive function on the disc such that its boundary value $\Omega|_\partial$ satisfies

$$\Omega|_\partial \gamma = \psi_*^{-1}\gamma \tag{5.24}$$

that is, so that the boundary condition on the metric is preserved. This action is well-defined, for suppose one chooses different extensions Ψ' and Ω', associated with the same ψ, then

$$\begin{aligned}\psi[h] &= [\Psi'_*\Omega' h] \\ &= [(\Psi' \circ \Psi^{-1})_*\Psi_*(\Omega'\Omega^{-1})\Omega h] \\ &= [(\Psi' \circ \Psi^{-1})_*((\Omega'\Omega^{-1}) \circ \Psi^{-1})\Psi_*\Omega h] \\ &= [\Psi_*\Omega h]\end{aligned} \tag{5.25}$$

where from the second to the third line we used the relation $\Psi_*\Omega T = (\Psi_*\Omega)\Psi_*T = (\Omega \circ \Psi^{-1})\Psi_*T$, for any diffeomorphism Ψ, multiplicative scalar Ω and tensor T; and from the third to the fourth line we used that $\Psi'\circ\Psi^{-1}$ and $(\Omega'\Omega^{-1})\circ\Psi^{-1}$ are trivial at the boundary, due to the boundary conditions, so their action is within the $\mathrm{Con}(D^*)$ classes. Thus the action depends only on ψ, not on the extensions. It also does not depend on the metric representative, i.e., $\psi[h] = \psi[\Phi_*\Lambda h]$ if $(\Phi, \Lambda) \in \mathrm{Con}(D^*)$, stabilising that the action is indeed well-defined. Finally, note that this action is transitive on $\mathcal{Q}$, due to Riemann mapping theorem, so that $\mathcal{Q}$ is a homogeneous space for $\mathit{Diff}^+(S^1)$. Consequently, $\mathcal{Q} = \mathit{Diff}^+(S^1)/H$ for some little group $H \in \mathit{Diff}^+(S^1)$.

The little group is the subgroup of $\mathit{Diff}^+(S^1)$ that leaves any particular metric h on the disc invariant.[22] For convenience, and without loss of generality, choose the metric $\bar{h}$ corresponding to the Euclidean round disc, described in polar coordinates as

$$\bar{h} = dr^2 + r^2 d\theta^2 \tag{5.26}$$

where $r \in [0, \ell/2\pi]$ and $\theta = [0, 2\pi)$. If $\psi \in H$, then it must extend to a *conformal isometry* of the disc,

$$\mathrm{ConIso}(D) := \mathrm{ConIso}(\bar{h}) := \{(\Upsilon, \Theta) \in \mathrm{Con}(D),\ \Upsilon_* \Theta \bar{h} = \bar{h}\} \tag{5.27}$$

and any conformal isometry of $\bar{h}$ must correspond to a $\upsilon = \Upsilon|_\partial \in H$. Therefore, we conclude that H is isomorphic to $\mathrm{ConIso}(D)$, so that

$$\mathcal{Q} = \mathrm{Riem}(D; \gamma)/\mathrm{Con}(D^*) = \mathit{Diff}^+(S^1)/\mathrm{ConIso}(D) \tag{5.28}$$

where $\mathrm{ConIso}(D)$ is seen as a subgroup of $\mathit{Diff}^+(S^1)$, defined by restricting $(\Upsilon, \Theta) \mapsto \Upsilon|_\partial$, and the quotient is from the right, i.e., $\psi \sim \psi \circ \upsilon$ iff $\upsilon \in \mathrm{ConIso}(D)$.

The group of conformal isometries of the round Euclidean disc is known to be $PSL(2, \mathbb{R})$ [71]. Here we present a quick review of the proof. We assume for simplicity that $\ell = 2\pi$, so the disc has unit radius — the group of conformal isometries is

[22] The group of conformal isometries for any two metrics on the disc are homomorphic: if (Υ, Θ) is a conformal isometry for the metric h, then $(\Psi, \Omega)(\Upsilon, \Theta)(\Psi, \Omega)^{-1}$ is a conformal isometry for $h' := \Psi_* \Omega h$.

clearly insensitive to the boundary length. The idea is to rephrase the problem in the language of complex analysis: the unit Euclidean disc can be naturally identified with the unit complex disc

$$\mathbb{D} = \{z \in \mathbb{C},\ |z| \leq 1\} \tag{5.29}$$

with the standard flat metric on $\mathbb{C}$. An (orientation-preserving) conformal isometry of the disc then translates into a biholomorphic map $f : \mathbb{D} \to \mathbb{D}$, as can be seen in App. B, particularly equation (B.9). Any such f, which we call a *conformal automorphism* of $\mathbb{D}$, must map the interior of $\mathbb{D}$ onto itself, $|f(z)| < 1 \iff |z| < 1$, and its boundary $\partial\mathbb{D}$ onto itself, $|f(z)| = 1 \iff |z| = 1$. Consider the Mobius transformation given by

$$M(z) = \frac{z - b}{1 - \bar{b}z} \tag{5.30}$$

where $b := f(0) \in \mathbb{D}$, with $|b| < 1$. Notice that this is a conformal automorphism of $\mathbb{D}$, as $|M(z)| = 1 \iff |z| = 1$ and $M(0) = -b \in \mathit{int}(\mathbb{D})$. Since the space of conformal transformations form a group (with the multiplication being the map composition), we have that the map $F = M \circ f$ is a conformal automorphism of D. Moreover, it satisfies $F(0) = 0$. By Schwarz lemma, applied to both F and F^{-1}, we must have

$$F(z) = e^{\imath\beta} z \tag{5.31}$$

for some $\beta \in \mathbb{R}$. Hence,

$$f(z) = M^{-1} \circ F(z) = M^{-1}(e^{\imath\beta} z) = e^{\imath\beta}\,\frac{z + e^{-\imath\beta} b}{1 + e^{\imath\beta}\bar{b} z} = e^{\imath\beta}\,\frac{z - a}{1 - \bar{a} z} \tag{5.32}$$

where $a := -e^{-\imath\beta} b$ is just another number in $\mathbb{D}$ satisfying $|a| < 1$. This is, therefore, the most general form of a conformal automorphism of the unit disc, characterizing ConIso(D).

Note that ConIso(D) forms a 3-dimensional group with topology $S^1 \times \mathbb{R}^2$. In fact, this group is precisely $PSL(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})/Z_2$, the projective special linear group in two real dimensions, i.e., consisting of 2×2 real matrices with unit determinant modded by the center $\{I, -I\}$. The way in which $PSL(2, \mathbb{R})$ appears most explicitly is by studying the group of conformal automorphisms of the complex upper plane

$$\mathbb{H} = \{z \in \mathbb{C},\ Im(z) \geq 0\} \tag{5.33}$$

The group of automorphisms consists of transformations

$$z \mapsto \frac{az + b}{cz + d} \tag{5.34}$$

where $a, b, c, d \in \mathbb{R}$ and $ad - bc > 0$. But notice that an overall scaling of $(a, b, c, d) \mapsto \lambda(a, b, c, d)$ does not affect a transformation, so we can restrict to $ad - bc = 1$; however, this leaves a residual $(a, b, c, d) \mapsto -(a, b, c, d)$ that needs to be modded out.

This set of transformations thus corresponds to $PSL(2,\mathbb{R})$. Since $\mathbb{H}$ is conformally equivalent to $\mathbb{D}$ (i.e., they are related by a biholomorphic map), ConIso(D) is homomorphic to $PSL(2,\mathbb{R})$.

The group of conformal automorphisms of the disc can be seen as a subgroup of $\mathit{Diff}^+(S^1)$ by restricting its action to the boundary. If a point at coordinate θ in $\partial\Sigma$ is represented by $z = e^{\imath\theta}$ in D, then

$$f(e^{\imath\theta}) = e^{\imath\beta}\,\frac{e^{\imath\theta} - \rho e^{\imath\alpha}}{\rho e^{\imath(\theta-\alpha)} - 1} = e^{\imath(\theta+\beta+2\gamma)} \tag{5.35}$$

where $a = \rho e^{\imath\alpha}$, with $\rho \in [0,1)$ and $\alpha \in [0,2\pi)$, and

$$\gamma = \arcsin\left[\frac{\rho\sin(\theta-\alpha)}{\sqrt{1+\rho^2-2\rho\cos(\theta-\alpha)}}\right] \tag{5.36}$$

with $\gamma \in (-\pi/2, \pi/2]$. Hence, the associated map $\upsilon = \Upsilon|_\partial$ on $\partial\Sigma$, defined from $e^{i\upsilon(\theta)} = f(e^{i\theta})$, is

$$\upsilon(\theta) = \theta + \beta + 2\arcsin\left[\frac{\rho\sin(\theta-\alpha)}{\sqrt{1+\rho^2-2\rho\cos(\theta-\alpha)}}\right] \tag{5.37}$$

which sits inside $\mathit{Diff}^+(S^1)$.

6 Reduction via conformal coordinates

In this section we consider another approach to the phase space reduction. This is based on a suitable "coordinate change" motivated by the previous result: instead of parametrizing the configuration space by spatial metrics, we parametrize it by conformal maps, as we will explain. This approach serves as a confirmation for the previous results and it also has the advantage of providing an explicit map between the physical, gauge-invariant degrees of freedom and the more concrete (but redundant) geometrical variables such as the spatial metric and extrinsic curvature.

6.1 Conformal coordinates

By virtue of the uniformization theorem, any Riemannian metric h_{ab} on $\Sigma \sim D$ can be obtained from a reference metric $\bar{h}_{ab}$ via some conformal transformation. That is, there exists a (orientation-preserving) diffeomorphism $\Psi : D \to D$ and a positive scalar $\Phi : D \to \mathbb{R}^+$ such that

$$h_{ab} = \Psi_* \Phi \bar{h}_{ab} \tag{6.1}$$

Because of the boundary condition on h, $h|_{\partial D} = \gamma$, the boundary value of Φ is determined from the boundary action of Ψ, $\psi := \Psi|_{\partial D}$,

$$\Phi \bar{h}\big|_{\partial D} = \psi_*^{-1} \gamma \tag{6.2}$$

In this way, we can use conformal maps $(\Psi, \Phi) \in \mathrm{Diff}^+(D) \times_\gamma C^\infty(D, \mathbb{R}^+)$, where $\times_\gamma$ indicates that the boundary condition (6.2) on Φ is satisfied,[23] as "coordinates" for the configuration space $\mathcal{Q} \sim \mathrm{Riem}(D, \gamma)$. Note that the map (Ψ, Φ) transforming $\bar{h}$ into h is not unique, since we can always compose it (on the right) with a conformal isometry (Ψ_0, Φ_0) of $\bar{h}$, that is, if $\bar{h}_{ab} = \Psi_{0*}\Phi_0\bar{h}_{ab}$, then $(\Psi', \Phi') := (\Psi, \Phi) \circ (\Psi_0, \Phi_0)$ also maps $\bar{h}$ into h. But this is not a problem since these "conformal coordinates" still cover the whole configuration space, and this non-uniqueness only amounts to additional gauge ambiguities being introduced in the description, which will all be removed in the end.

To cover the phase space we need also the momentum "coordinates" associated with the extrinsic curvature. As our choice of time gauge has eliminated the trace-part of K_{ab}, we must consider only its traceless part, σ^{ab}. It is convenient to use the conformal map (Ψ, Φ) to transform σ^{ab} to the reference disc D. So we define the transformed traceless extrinsic curvature, $\bar{\sigma}$, as

$$\bar{\sigma}^{ab} := \Phi^{-s}\Psi_*^{-1}\sigma^{ab} \tag{6.3}$$

where s is a power which will be chosen so as to simplify the momentum constraint. This can be equivalently written as

$$\sigma^{ab} = \Psi_*\Phi^s\bar{\sigma}^{ab} \tag{6.4}$$

[23]More precisely, $\mathrm{Diff}^+(D) \times_\gamma C^\infty(D, \mathbb{R}^+) := \{(\Psi, \Phi) \in \mathrm{Diff}^+(D) \times C^\infty(D, \mathbb{R}^+);\, \Phi\bar{h}\big|_{\partial D} = \psi_*^{-1}\gamma\}$.

evidencing the isomorphism between σ's and $\bar{\sigma}$'s defined by the conformal map (Ψ, Φ). In this way, we see that the triplet $(\Psi, \Phi, \bar{\sigma})$ can be used as "coordinates" for the (unconstrained) phase space. This can be seen as an *enlargement* of the phase space from $\mathcal{P} = \mathrm{Riem}(D, \gamma) \times \mathrm{Sym}(D, (2,0))$ to

$$\bar{\mathcal{P}} := \mathrm{Diff}^{+}(D) \times_{\gamma} C^{\infty}(D, \mathbb{R}^{+}) \times \mathrm{Sym}(D, (2,0)) \tag{6.5}$$

where $\times_{\gamma}$ refers to condition (6.2). As mentioned before, this enlargement consists of extra gauge directions being introduced, which will be eventually removed.

6.2 Imposing the constraints

Now we impose the constraints. First, consider the momentum constraint, (3.18), which in the CMC gauge becomes

$$\nabla_{a}\sigma^{ab} = 0 \tag{6.6}$$

that is, σ must be *transverse* with respect to h. Now we want to investigate how this condition translates when transformed to D, i.e., when expressed in terms of $\bar{h}$ and $\bar{\sigma}$. Since ∇ is covariantly constructed from h, equation (6.6) transforms nicely under diffeomorphisms,

$$\Psi_{*}^{-1}\left(\nabla_{a}^{(h)}\sigma^{ab}\right) = \nabla_{a}^{(\Psi_{*}^{-1}h)}\left(\Psi_{*}^{-1}\sigma^{ab}\right) = 0 \tag{6.7}$$

so we have

$$\nabla_a^{(\Phi\bar{h})}\left(\Phi^s\bar{\sigma}^{ab}\right) = 0 \tag{6.8}$$

The left-hand side corresponds to a Weyl transformation, which yields

$$\nabla_a^{(\Phi\bar{h})}\left(\Phi^s\bar{\sigma}^{ab}\right) = \Phi^s\bar{\nabla}_a\bar{\sigma}^{ab} + (s+2)\Phi^{s-1}\bar{\sigma}^{ab}\bar{\nabla}_a\Phi - \frac{1}{2}\Phi^{s-1}\bar{h}_{cd}\bar{\sigma}^{cd}\bar{h}^{ab}\bar{\nabla}_a\Phi = 0 \tag{6.9}$$

where $\bar{\nabla}$ is the covariant derivative associated with $\bar{h}$. The last term vanishes because the tracelessness of σ (with respect to h) implies that $\bar{\sigma}$ is traceless (with respect to $\bar{h}$). Thus we see that the choice $s = -2$ is particularly convenient as it makes condition (6.6) equivalent to

$$\bar{\nabla}_a\bar{\sigma}^{ab} = 0 \tag{6.10}$$

that is, $\bar{\sigma}$ must be transverse with respect to $\bar{h}$. Second, consider the Hamiltonian constraint (3.17), which in the CMC slice with $K = -\tau$ becomes

$$\sigma^{ab}\sigma_{ab} - R^{(h)} - \chi = 0 \tag{6.11}$$

where $\chi = -2\Lambda + \frac{1}{2}\tau^2$ and $R^{(h)}$ is the Ricci scalar associated with h. As before we apply Ψ_*^{-1} to this expression to get

$$(\Psi_*^{-1}h_{ac})(\Psi_*^{-1}h_{bd})(\Psi_*^{-1}\sigma^{ab})(\Psi_*^{-1}\sigma^{cd}) - R^{(\Psi_*^{-1}h)} - \chi = 0 \tag{6.12}$$

that is,

$$(\Phi\bar{h}_{ac})(\Phi\bar{h}_{bd})(\Phi^{-2}\bar{\sigma}^{ab})(\Phi^{-2}\bar{\sigma}^{cd}) - R^{(\Phi\bar{h})} - \chi = 0 \tag{6.13}$$

which yields

$$\bar{\nabla}^2\lambda - \bar{R} + e^{-\lambda}\bar{\sigma}^{ab}\bar{\sigma}_{ab} - e^{\lambda}\chi = 0 \tag{6.14}$$

where $\lambda = \log\Phi$, $\bar{R}$ is the Ricci scalar associated with $\bar{h}$ and $\bar{\sigma}^{ab}\bar{\sigma}_{ab}$ is contracted using $\bar{h}$. For convenience, we can choose $\bar{h}$ to be the metric on a flat unit disc

$$\bar{h} = dr^2 + r^2 d\theta^2 \tag{6.15}$$

where $(r, \theta) \in [0, 1] \times [0, 2\pi)$ are the usual polar coordinates on the disc. In this way, $\bar{R} = 0$.

To sum up, we are considering a change of coordinates on the unconstrained phase space from (h_{ab}, σ^{ab}) to $(\Psi, \lambda, \bar{\sigma}^{ab})$ defined by

$$h_{ab} = \Psi_* e^{\lambda}\bar{h}_{ab} \tag{6.16}$$

$$\sigma^{ab} = \Psi_* e^{-2\lambda}\bar{\sigma}^{ab} \tag{6.17}$$

where $\bar{h}_{ab}$ is the Euclidean metric on the reference disc D and $\bar{\sigma}^{ab}$ is a traceless (with respect to $\bar{h}$) symmetric tensor on D. Because of the boundary condition on h, $h|_{\partial\Sigma} = \gamma$, the boundary value of λ is determined from the boundary action of Ψ,

$\psi := \Psi|_{\partial D}$,

$$e^{\lambda}\bar{h}\Big|_{\partial D} = \psi_*^{-1}\gamma \tag{6.18}$$

The constraint surface $\mathcal{S}$ is determined by imposing condition (6.10) on $\bar{\sigma}$ and taking λ to satisfy

$$\bar{\nabla}^2\lambda + e^{-\lambda}\bar{\sigma}^{ab}\bar{\sigma}_{ab} - e^{\lambda}\chi = 0 \tag{6.19}$$

The arguments of the previous sections imply that (6.19) has unique solution depending on $\bar{\sigma}$ (via the $\bar{\sigma}^{ab}\bar{\sigma}_{ab}$ term) and on Ψ (via the boundary conditions). Therefore, $\mathcal{S}$ can be "covered" with coordinates $(\Psi, \bar{\sigma}^{ab})$, where $\bar{\sigma}$ is traceless and transverse with respect to $\bar{h}$. More precisely, the map $(\Psi, \bar{\sigma}^{ab}) \mapsto (h_{ab}, \sigma^{ab})$ is a projection map from $\text{Diff}^+(D) \times \text{Sym}(D, (2,0); TT[\bar{h}])$, where $TT[\bar{h}]$ means "traceless and transverse with respect to $\bar{h}$", onto $\mathcal{S}$. Note that we can think of

$$\bar{\mathcal{S}} := \text{Diff}^+(D) \times \text{Sym}(D, (2,0); TT[\bar{h}]) \tag{6.20}$$

as an *enlargement* of $\mathcal{S}$, where additional gauge ambiguities have been introduced. In fact, it is equivalent to think that we are first enlarging $\mathcal{P}$ into $\bar{\mathcal{P}}$ and then applying the constraints to get $\bar{\mathcal{S}}$, or to think that we are first applying the constraints on $\mathcal{P}$ to get $\mathcal{S}$ and then enlarging it into $\bar{\mathcal{S}}$. For concreteness, let us call the projection from $\bar{\mathcal{S}}$ to $\mathcal{S}$ by $T : \bar{\mathcal{S}} \to \mathcal{S}$, so that

$$T(\Psi, \bar{\sigma}^{ab}) = (h_{ab}, \sigma^{ab}) = (\Psi_* e^{\lambda}\bar{h}_{ab}, \Psi_* e^{-2\lambda}\bar{\sigma}^{ab}) \tag{6.21}$$

where λ satisfies (6.19).

Now we want to pull-back the (pre)symplectic form ω on $\mathcal{S}$ to $\bar{\mathcal{S}}$,

$$\bar{\omega} = T^* \omega \tag{6.22}$$

which corresponds to "writing ω in $(\Psi, \bar{\sigma})$ coordinates". The tangent vector $\bar{\eta}$ at a point $(\Psi, \bar{\sigma})$ of $\bar{\mathcal{S}}$ can be expressed as $(X, \alpha) \in T_\Psi \mathrm{Diff}^+(D) \oplus T_{\bar{\sigma}}[\mathrm{Sym}(D, (2,0); TT[\bar{h}])]$. Since $\mathrm{Sym}(D, (2,0); TT[\bar{h}])$ is a vector space, it can be naturally identified with its tangent space, so that $T_{\bar{\sigma}}[\mathrm{Sym}(D, (2,0); TT[\bar{h}])] \sim \mathrm{Sym}(D, (2,0); TT[\bar{h}])$. Informally speaking, X can be seen as the vector tangent to a one-parameter family of diffeomorphisms, $t \mapsto \Psi_t$. To be more precise, we can use the group structure of $\mathrm{Diff}^+(D)$ to left-translate X to the identity I, and note that the tangent space to I can be naturally identified with vector fields on D (satisfying the condition that the vectors at the boundary are tangent to it, since automorphisms of D must map ∂D into itself). Introducing some notation, let us call

$$\bar{X} := l_{\Psi^{-1}*} X \tag{6.23}$$

where $l_\Psi(\Psi') = \Psi \circ \Psi'$ is the left multiplication on $\mathrm{Diff}^+(D)$. As $\bar{X} \in T_I \mathrm{Diff}^+(D)$, let us denote the associated vector field on D by $\xi \in \mathrm{Vect}_0(D)$, where $\mathrm{Vect}_0(D)$ is the space of vector fields on D satisfying the parallel boundary condition $\xi|_{\partial D} \in T(\partial D)$. In this way, we can define X as being the vector tangent to the curve $t \mapsto \Psi \circ \Gamma_t$ at

$t = 0$, where $\Gamma_t = \mathrm{Exp}(t\xi)$ is the diffeomorphism corresponding to flowing along the integral curves of ξ for a parameter t. Thus, we have the identification

$$T_{(\Psi,\bar{\sigma})}\mathcal{S} \sim \mathrm{Vect}_0(D) \oplus \mathrm{Sym}(D,(2,0);TT[\bar{h}]) \tag{6.24}$$

which allows us to express $\bar{\eta}$ in the form

$$\bar{\eta} = (\xi^a, \alpha^{ab}) \tag{6.25}$$

where $\xi \in \mathrm{Vect}_0(D)$ and $\alpha \in \mathrm{Sym}(D,(2,0);TT[\bar{h}])$. If $\bar{\eta}$ and $\bar{\eta}'$ are vectors tangent to $\bar{\mathcal{S}}$ at some point $(\Psi, \bar{\sigma})$, and $\eta := T_*\bar{\eta}$ and $\eta' := T_*\bar{\eta}'$ are the corresponding pushed vector to $\mathcal{S}$, then

$$\bar{\omega}(\bar{\eta}, \bar{\eta}') = \omega(\eta, \eta') \tag{6.26}$$

Since ω is the restriction to $\mathcal{S}$ of Ω, given in (3.11), and η and η' are tangent to $\mathcal{S}$, we also have

$$\bar{\omega}(\bar{\eta}, \bar{\eta}') = \Omega(\eta, \eta') \tag{6.27}$$

Note that Ω involves the conjugate momentum π^{ab}, which from definition (3.12) and relation (3.16) can be written as

$$\pi^{ab} = \sqrt{\det(h)}\left(\sigma^{ab} + \frac{1}{2}\tau h^{ab}\right) \tag{6.28}$$

Taking the exterior derivative (on the unconstrained phase space $\mathcal{P}$) gives

$$\delta\pi^{ab} = \sqrt{\det(h)}\left[\delta\sigma^{ab} + \frac{1}{2}\left(\sigma^{ab}h^{cd} - \tau h^{ac}h^{bd} + \frac{1}{2}\tau h^{ab}h^{cd}\right)\delta h_{cd}\right] \tag{6.29}$$

and replacing this on Ω yields

$$\Omega = \int_D d^2x\, \delta\pi^{ab} \wedge \delta h_{ab} = \int_D \vartheta_h \left(\delta\sigma^{ab} + \frac{1}{2}\sigma^{ab}h^{cd}\delta h_{cd}\right) \wedge \delta h_{ab} \tag{6.30}$$

where

$$\vartheta_h = d^2x\sqrt{\det(h)} \tag{6.31}$$

is the natural volume form associated with h.[24]

In order to evaluate $\Omega(\eta, \eta')$, we compute the "variations" $\delta h_{ab}(\eta)$ and $\delta\sigma^{ab}(\eta)$ for a pushed vector $\eta = T_*\bar{\eta}$. First note that since $\mathrm{Riem}(D;\gamma)$ is a subspace of $\mathrm{Sym}(D,(0,2))$, we can think of a "variation of h" as a difference between nearby metrics, so that $\delta h_{ab}(\eta) \in \mathrm{Sym}(D,(0,2))$.[25] Similarly, $\mathrm{Sym}(D,(2,0))$ is a vector space, so that a "variation of σ" can also be thought of as a difference between nearby tensors and $\delta\sigma^{ab}(\eta) \in \mathrm{Sym}(D,(2,0))$. Thus, for $\bar{\eta} = (\xi, \alpha) \in T_{(\Psi,\bar{\sigma})}\bar{\mathcal{S}}$, $\delta h_{ab}(\eta)$

[24] In a covariant language, given an orientable manifold $\mathcal{M}$ with orientation n-form ϑ, the natural volume form ϑ_h associated with a metric h is defined as $\vartheta_h = w\vartheta$, where the scalar $w > 0$ is such that $\vartheta_h(e_1, e_2, \ldots e_n) = 1$ for any (oriented) orthonormal basis $\{e_1, e_2, \ldots e_n\}$.

[25] To say this more precisely, we are regarding $h_{ab}(x)$ as a function from $\mathcal{S}$ into $T_x^{(2,0),\mathrm{sym}}D$, the space of symmetric $(2,0)$ tensors at $x \in D$, where $h_{ab}(x)$ takes a point $p \in \mathcal{S}$ and returns the corresponding value of the metric at $x \in D$. Since the target space is a vector space, $\delta h_{ab}(x)$ can be defined in the usual manner, i.e., given a vector $\eta \in T_p\mathcal{S}$ we define $\delta h_{ab}(x)(\eta) := \eta(h_{ab}(x)) = \frac{d}{dt}h_{ab}(x)[p_t]$, where p_t is a curve on $\mathcal{S}$ tangent to η.

and $\delta\sigma^{ab}(\eta)$ are given by

$$\delta h_{ab}(\eta) = \frac{d}{dt}\left[(\Psi \circ \Gamma_t)_* e^{\lambda_t} \bar{h}_{ab}\right] \tag{6.32}$$

$$\delta\sigma^{ab}(\eta) = \frac{d}{dt}\left[(\Psi \circ \Gamma_t)_* e^{-2\lambda_t} \bar{\sigma}_t^{ab}\right] \tag{6.33}$$

where the derivative is evaluated at $t = 0$, $\Gamma_t = \mathrm{Exp}(t\xi)$, $\sigma_t^{ab} = \alpha^{ab} t$ and λ_t is the solution of (6.19) associated with $(\Psi \circ \Gamma_t, \bar{\sigma}_t)$. By distributing the derivative we get

$$\delta h_{ab}(\eta) = \Psi_* \left[\frac{d}{dt}\Gamma_{t*}(e^{\lambda}\bar{h}_{ab}) + e^{\lambda}\bar{h}_{ab}\frac{d\lambda_t}{dt}\right] \tag{6.34}$$

$$\delta\sigma^{ab}(\eta) = \Psi_* \left[\frac{d}{dt}\Gamma_{t*}(e^{-2\lambda}\bar{\sigma}^{ab}) - 2e^{-2\lambda}\bar{\sigma}_t^{ab}\frac{d\lambda_t}{dt} + e^{-2\lambda}\frac{d\bar{\sigma}_t^{ab}}{dt}\right] \tag{6.35}$$

which yields

$$\delta h_{ab}(\eta) = \Psi_* \left[-\pounds_{\xi}(e^{\lambda}\bar{h}_{ab}) + \kappa e^{\lambda}\bar{h}_{ab}\right] \tag{6.36}$$

$$\delta\sigma^{ab}(\eta) = \Psi_* \left[-\pounds_{\xi}(e^{-2\lambda}\bar{\sigma}^{ab}) - 2\kappa e^{-2\lambda}\bar{\sigma}^{ab} + e^{-2\lambda}\alpha^{ab}\right] \tag{6.37}$$

where κ is defined as

$$\kappa(x; \Psi, \bar{\sigma}, \eta) := \left.\frac{d\lambda_t}{dt}\right|_{t=0} \tag{6.38}$$

which is a function of $x \in D$ depending implicitly on $(\Psi, \bar{\sigma})$ and η. Thus for the first term in (6.30) we obtain

$$\delta\sigma^{ab} \wedge \delta h_{ab}(\eta, \eta') = \Psi_* e^{-\lambda} \Big[\pounds_\xi \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} - 2(\pounds_\xi \lambda - \kappa)\bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} + \\ + (\pounds_{\xi'} \lambda - \kappa') \pounds_\xi \bar{\sigma}^{ab} \bar{h}_{ab} - \pounds_{\xi'} \bar{h}_{ab} \alpha^{ab} - \{(\xi, \alpha) \leftrightarrow (\xi', \alpha')\} \Big] \quad (6.39)$$

where $\kappa' = \kappa(x; \Psi, \bar{\sigma}, \eta')$ and $\{(\xi, \alpha) \leftrightarrow (\xi', \alpha')\}$ consists of the previous terms but with (ξ, α) exchanged with (ξ', α'). Also, for the second term in (6.30), we have

$$\frac{1}{2}\sigma^{ab} h^{cd} \delta h_{cd} \wedge \delta h_{ab}(\eta, \eta') = \Psi_* e^{-\lambda} \Big[\frac{1}{2} \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} \bar{h}^{cd} \pounds_\xi \bar{h}_{cd} + \\ + (\pounds_\xi \lambda - \kappa) \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} - \{(\xi, \alpha) \leftrightarrow (\xi', \alpha')\} \Big] \qquad (6.40)$$

Therefore,

$$\Omega(\eta, \eta') = \int_D \vartheta_{e^\lambda \bar{h}} e^{-\lambda} \left[\pounds_\xi \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} - \pounds_{\xi'} \bar{h}_{ab} \alpha^{ab} + \frac{1}{2} \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} \bar{h}^{cd} \pounds_\xi \bar{h}_{cd} - \{(\xi, \alpha) \leftrightarrow (\xi', \alpha')\} \right] \qquad (6.41)$$

where we have used that $\vartheta_h = \vartheta_{\Psi_* e^\lambda \bar{h}} = \Psi_* \vartheta_{e^\lambda \bar{h}}$ and that integrals are invariant under automorphisms, i.e., $\int A = \int \Psi_* A$. Note that the factor $(\pounds_\xi \lambda - \kappa)$ does not appear because it multiplies $\pounds_{\xi'}(\bar{\sigma}^{ab} \bar{h}_{ab})$, which vanishes due to the tracelessness of $\bar{\sigma}$. Now using that

$$\vartheta_{e^\lambda \bar{h}} = e^\lambda \vartheta_{\bar{h}} \qquad (6.42)$$

we get

$$\Omega(\eta,\eta') = \int_D \vartheta_{\bar{h}} \left[\pounds_\xi \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} - \pounds_{\xi'} \bar{h}_{ab} \alpha^{ab} + \frac{1}{2} \bar{\sigma}^{ab} \pounds_{\xi'} \bar{h}_{ab} \bar{h}^{cd} \pounds_\xi \bar{h}_{cd} - \{(\xi,\alpha) \leftrightarrow (\xi',\alpha')\} \right] \tag{6.43}$$

which is an integral defined entirely on the reference disc $(D;\bar{h})$. Writing the Lie derivatives in terms of covariant derivatives (with respect to $\bar{h}$), i.e., $\pounds_\xi \bar{h}_{ab} = 2\bar{\nabla}_{(a}\xi_{b)}$ and $\pounds_\xi \bar{\sigma}^{ab} = \xi^c \bar{\nabla}_c \bar{\sigma}^{ab} - 2\bar{\nabla}_c \xi^{(a} \bar{\sigma}^{b)c}$, we obtain

$$\begin{aligned} \bar{\omega}(\bar{\eta},\bar{\eta}') = \Omega(\eta,\eta') &= 2\int_D \vartheta_{\bar{h}} \bar{\nabla}_a \left[\left(\xi_b \alpha'^{ab} - \xi'_b \alpha^{ab} \right) - \left(\xi^c \bar{\nabla}_c \xi'_b - \xi'_c \bar{\nabla}_c \xi_b \right) \bar{\sigma}^{ab} \right] \\ &= 2\int_{\partial D} d\theta \, n_a \left[\left(\xi_b \alpha'^{ab} - \xi'_b \alpha^{ab} \right) - [\xi,\xi']_b \bar{\sigma}^{ab} \right] \end{aligned} \tag{6.44}$$

In the second line we have used Gauss's law, so n^a is a unit (outward-pointing) normal vector at $\partial D \sim S^1$ and $d\theta$ is the measure induced on ∂D. Since $\bar{h}$ was chosen to be the unit-radius metric on the disc, $d\theta$ is just the differential of the angle coordinate, $\theta \in [0, 2\pi)$.

Before fully appreciating the previous result, it is useful to understand better a quantity like $\int d\theta \, n_a \bar{\sigma}^{ab} \xi_b$, where n is normal and ξ is tangent to the boundary. In Cartesian coordinates $\{x, y\}$ on $(D, \bar{h})$, the symmetry and traceless of $\bar{\sigma}$ implies that its components have generic form

$$\bar{\sigma}^{\mu\nu} = \begin{pmatrix} v & u \\ u & -v \end{pmatrix} \tag{6.45}$$

for u, $v \in C^\infty(D, \mathbb{R})$. It is convenient to think of these components as parts of a complex function $f : D \to \mathbb{C}$ defined by

$$f(z) := u(z) + iv(z) \tag{6.46}$$

where $z = x+iy$, $u(z) := u(x, y)$ and similarly for v. In this notation, we immediately see that the condition that $\bar{\sigma}$ is transverse to $\bar{h}$, $\partial_\mu \bar{\sigma}^{\mu\nu} = 0$, translates into the condition that f is analytic,

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \end{aligned} \tag{6.47}$$

In particular this implies that f is completely determined on D from its value on the boundary ∂D. To see this note that u and v are harmonic functions, that is, they satisfy

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \tag{6.48}$$

Therefore, the general solution for u on D is

$$u = \sum_{n \geq 0} r^n \left(a_n \cos n\theta + b_n \sin n\theta \right) \tag{6.49}$$

where $\{r, \theta\}$ are the usual polar coordinates (defined by $x = r\cos\theta$ and $y = r\sin\theta$). It is clear that u is uniquely determined from its boundary value $\hat{u}(\theta) := u(r = 1, \theta)$.

Now note that

$$dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy = -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy \tag{6.50}$$

so that

$$v(x,y) = v(0,0) + \int_\gamma \left(-\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy\right) \tag{6.51}$$

where γ is any curve joining $(0,0)$ to $(x,y) \in D$. We see that v is completely determined from u, modulo the choice of its value v_0 at the origin,

$$v = v_0 + \sum_{n\geq 0} r^n \left(b_n \cos n\theta - a_n \sin n\theta\right) \tag{6.52}$$

Note that v_0 can be absorbed in b_0, so that the general solution for $\bar{\sigma}$ can be written as

$$\bar{\sigma}^{\mu\nu} = \sum_{n\geq 0} r^n \left[a_n \begin{pmatrix} \sin n\theta & \cos n\theta \\ \cos n\theta & -\sin n\theta \end{pmatrix} + b_n \begin{pmatrix} -\cos n\theta & \sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}\right] \tag{6.53}$$

for coefficients a_n and b_n. This expansion (6.53) will be taken as defining the coefficients a_n and b_n, which will be called the Fourier coefficients of $\bar{\sigma}$. Let n be the unit normal vector field and $t = \partial_\theta$ be the unit tangent vector field on the boundary, $\partial D \sim S^1$. That is, at the point $(x,y) = (\cos\theta, \sin\theta)$, we have $n = \cos\theta\partial_x + \sin\theta\partial_y$ and $t = \cos\theta\partial_x - \sin\theta\partial_y$. Then,

$$\bar{\sigma}^{ab} n_a t_b = u\cos 2\theta - v \sin 2\theta = \sum_{n\geq 0} \left[a_n \cos(n+2)\theta + b_n \sin(n+2)\theta\right] \tag{6.54}$$

If $\xi = f\partial_\theta$ is a tangent vector field on the boundary, where $f \in C^\infty(S^1, \mathbb{R})$, we have

$$\int d\theta\, \bar{\sigma}^{ab} n_a \xi_b = \int d\theta \sum_{n\geq 0} \left[a_n \cos(n+2)\theta + b_n \sin(n+2)\theta\right] f(\theta) \tag{6.55}$$

From the Fourier coefficients of $\bar{\sigma}$, we define the quadratic form $\mathring{\sigma}$ on the boundary by

$$\mathring{\sigma}(\theta) := -2 \sum_{n\geq 0} \left[a_n \cos(n+2)\theta + b_n \sin(n+2)\theta\right] d\theta^2 \tag{6.56}$$

This provides a convenient identification between $\bar{\sigma}^{ab}$ (a symmetric, traceless and transverse tensor on D) and $\mathring{\sigma}$ (a quadratic form on $\partial D \sim S^1$ "missing" the Fourier modes 1, $\sin\theta$ and $\cos\theta$). It is also convenient to introduce some notation for this space of $\mathring{\sigma}$'s. The Lie algebra of *Diff*$^+(S^1)$, denoted $\mathfrak{diff}(S^1)$, is naturally represented by vector fields $\hat{\xi}$ on S^1. We can write it as $\hat{\xi} = f\partial_\theta$, where $f \in C^\infty(S^1, \mathbb{R})$. The dual Lie algebra, denoted $\mathfrak{diff}^*(S^1)$, is naturally represented by quadratic forms α on S^1. We can write it as $\alpha = a(\theta)d\theta^2$, where $a \in C^\infty(S^1, \mathbb{R})$. The reason for representing dual vectors in this way is because it leads to the natural pairing

$$\alpha(\hat{\xi}) := \int a d\theta^2 (f\partial_\theta) = \int d\theta\, a(\theta) f(\theta) \tag{6.57}$$

and this will simplify certain formulas (especially in the part about coadjoint orbits). Now observe that the space of $\mathring{\sigma}$'s, as defined in (6.56), is the subspace $\mathring{\mathfrak{diff}}^*(S^1)$ of $\mathfrak{diff}^*(S^1)$ "missing" the Fourier modes 1, $\sin\theta$ and $\cos\theta$. Another way to characterize $\mathring{\mathfrak{diff}}^*(S^1)$ is to notice that its elements annihilate the generators

$\hat{\xi} = \partial_\theta$, $\sin\theta\,\partial_\theta$, $\cos\theta\,\partial_\theta$ of $\mathfrak{psl}(2,\mathbb{R}) \subset \mathfrak{diff}(S^1)$, that is,

$$\mathring{\mathfrak{diff}}^*(S^1) := \{\mathring{\sigma} \in \mathfrak{diff}^*(S^1);\text{ where } \mathring{\sigma}(\hat{\xi}) = 0 \text{ for all } \hat{\xi} \in \mathfrak{psl}(2,\mathbb{R}) \subset \mathfrak{diff}(S^1)\} \quad (6.58)$$

To sum up, we have identified

$$\mathrm{Sym}(D,(2,0);TT[\bar{h}]) = \mathring{\mathfrak{diff}}^*(S^1) \quad (6.59)$$

in a natural way. In this language, the quantity in (6.55) can be expressed as

$$\int d\theta\,\bar{\sigma}^{ab} n_a \xi_b = -\frac{1}{2}\mathring{\sigma}(\hat{\xi}) \quad (6.60)$$

where $\hat{\xi} \in \mathfrak{diff}(S^1)$ is simply the restriction of ξ to $\partial D \sim S^1$, which is well-defined since ξ is tangent to ∂D. Correspondingly, the symplectic form in (6.44) takes the form

$$\bar{\omega}(\bar{\eta},\bar{\eta}') = \mathring{\alpha}(\hat{\xi}') - \mathring{\alpha}'(\hat{\xi}) + \mathring{\sigma}([\hat{\xi},\hat{\xi}']) \quad (6.61)$$

where $\mathring{\alpha} \in \mathring{\mathfrak{diff}}^*(S^1)$ is related to α^{ab} in the same way as in (6.56), and analogously for $\hat{\alpha}'$ and α'^{ab}. Note that the Lie brackets is independent on taking the restriction to ∂D before or after computing the bracket, i.e., $[\hat{\xi},\hat{\xi}'] = \widehat{[\xi,\xi']}$.

6.3 Removing the bulk diffeomorphisms

The expression above for the symplectic form on $\bar{\mathcal{S}} = \text{Diff}^+(D)\times\text{Sym}(D,(2,0);TT[\bar{h}])$ clearly reveals a "huge" part of the gauge ambiguities. That is, bulk diffeomorphisms acting trivially on the boundary are pure gauge transformations. In particular, this means that any $\bar{\eta} = (\xi, \alpha)$ where ξ *vanishes at the boundary* corresponds to a degenerate direction of $\bar{\omega}$. Therefore, any two points $(\Psi, \bar{\sigma})$ and $(\Psi', \bar{\sigma})$ in $\bar{\mathcal{S}}$ such that the diffeomorphisms Ψ and Ψ' have the same boundary action, $\Psi|_{\partial D} = \Psi'|_{\partial D}$, correspond to the same physical state. Thus, we have the following reduction

$$\begin{aligned} \text{Diff}^+(D)\times\text{Sym}(D,(2,0);TT[\bar{h}]) &\to \mathit{Diff}^+(S^1)\times\mathring{\mathfrak{diff}}^*(S^1) \\ (\Psi, \bar{\sigma}^{ab}) &\mapsto (\psi, \mathring{\sigma}) \end{aligned} \tag{6.62}$$

where $\psi \in \mathit{Diff}^+(S^1)$,

$$\psi := \Psi\Big|_{\partial D} \tag{6.63}$$

is the boundary action of Ψ. Note that the quotient is really acting only on this diffeomorphism factor, $\Psi \mapsto \psi$, while the association $\bar{\sigma}^{ab} \mapsto \mathring{\sigma}$ is just the isomorphism discussed above. For concreteness, let us refer to this phase space as

$$\hat{\mathcal{S}} := \mathit{Diff}^+(S^1)\times\mathring{\mathfrak{diff}}^*(S^1) \tag{6.64}$$

and to the corresponding quotient map as $R : \bar{\mathcal{S}} \to \hat{\mathcal{S}}$. The symplectic form on $\hat{\mathcal{S}}$, denoted by $\hat{\omega}$, must satisfy $\bar{\omega} = R^* \hat{\omega}$. First note that the push-forward of a vector $\bar{\eta} = (\xi, \alpha^{ab})$ in $\bar{\mathcal{S}}$ is given simply by[26]

$$R_*(\xi, \alpha^{ab}) = (\hat{\xi}, \mathring{\alpha}) \tag{6.65}$$

where $\hat{\xi}$ is, as before, the restriction of ξ to the boundary. Thus, one can easily verify that

$$\hat{\omega}(\hat{\eta}, \hat{\eta}') = \mathring{\alpha}(\hat{\xi}') - \mathring{\alpha}'(\hat{\xi}) + \mathring{\sigma}([\hat{\xi}, \hat{\xi}']) \tag{6.66}$$

defines the desired symplectic form $\hat{\omega}$ on $\hat{\mathcal{S}}$, where $\hat{\eta} = (\hat{\xi}, \mathring{\alpha})$ and $\hat{\eta}' = (\hat{\xi}', \mathring{\alpha}')$ are tangent vectors at $(\psi, \mathring{\sigma})$.

We must investigate if there are remaining gauge directions to be eliminated, that is, if $\hat{\omega}$ has is degenerate (and thus possess null directions). Suppose that $\hat{\eta} = (\hat{\xi}, \hat{\alpha})$ is a null direction of $\hat{\omega}$, $\imath_{\hat{\eta}} \hat{\omega} = 0$, that is, $\hat{\omega}(\hat{\eta}, \hat{\eta}') = 0$ for all vectors $\hat{\eta}'$. First let $\hat{\eta}' = (0, \mathring{\alpha}')$, for a generic $\mathring{\alpha}' \in \mathring{\mathfrak{diff}}^*(S^1)$. This implies that

$$\hat{\omega}(\bar{\eta}, \bar{\eta}') = -\mathring{\alpha}'(\hat{\xi}) = 0 \tag{6.67}$$

But from the definition of $\mathring{\mathfrak{diff}}^*(S^1)$, (6.58), we have that only the $\mathfrak{psl}(2, \mathbb{R})$ subset of $\mathfrak{diff}(S^1)$ is annihilated by all $\mathring{\alpha}'$. Therefore, if $\hat{\eta}$ is a null direction, we must have

[26]The tangent space to a point $(\psi, \mathring{\sigma}) \in \hat{\mathcal{S}}$ is being characterized in a way analogous to how we characterized the tangent space for $\bar{\mathcal{S}}$, that is, $T_{(\psi, \mathring{\sigma})} \hat{\mathcal{S}} \sim \mathfrak{diff}(S^1) \oplus \mathring{\mathfrak{diff}}^*(S^1)$, so a generic vector is represented by a pair $\hat{\eta} = (\hat{\xi}, \hat{\alpha})$, where $\hat{\xi}$ is a vector field on S^1 and $\hat{\alpha}$ is an element of $\mathring{\mathfrak{diff}}^*(S^1)$.

$\hat{\xi} \in \partial_\theta \oplus \sin\theta\,\partial_\theta \oplus \cos\theta\,\partial_\theta$. Now consider $\hat{\eta}' = (\hat{\xi}', 0)$, for a generic $\hat{\xi}' \in \mathfrak{diff}(S^1)$. This implies that

$$\hat{\omega}(\bar{\eta}, \bar{\eta}') = \mathring{\alpha}(\hat{\xi}') + \mathring{\sigma}([\hat{\xi}, \hat{\xi}']) = 0 \tag{6.68}$$

Let $\mathring{\alpha} = ad\theta^2$, $\hat{\xi} = f\partial_\theta$, $\hat{\xi}' = f'\partial_\theta$ and, with a slight abuse of notation, $\mathring{\sigma} = \sigma d\theta^2$. Then, according to the pairing defined in (6.57), we have

$$\hat{\omega}(\bar{\eta}, \bar{\eta}') = \int d\theta \left(af' + \sigma(f\partial_\theta f' - f'\partial_\theta f)\right) = \int d\theta\, f' \left(a - f\partial_\theta\sigma - 2\sigma\partial_\theta f\right) \tag{6.69}$$

where we used $[f\partial_\theta, f'\partial_\theta] = (f\partial_\theta f' - f'\partial_\theta f)\partial_\theta$ and, in the last equality, we integrated by parts so as to factorize f'. Note that since f' is arbitrary, we conclude that $\hat{\eta}$ is null only if

$$a = f\partial_\theta\sigma + 2\sigma\partial_\theta f \tag{6.70}$$

Therefore, we have showed that $\hat{\omega}$ has exactly three degenerate directions at a generic point $(\hat{\psi}, \sigma d\theta^2)$,

$$\begin{aligned}
\hat{\eta}_1 &= (\partial_\theta, \partial_\theta\sigma d\theta^2) \\
\hat{\eta}_2 &= (\sin\theta\,\partial_\theta, (\sin\theta\,\partial_\theta\sigma + 2\cos\theta\,\sigma)d\theta^2) \\
\hat{\eta}_3 &= (\cos\theta\,\partial_\theta, (\cos\theta\,\partial_\theta\sigma - 2\sin\theta\,\sigma)d\theta^2)
\end{aligned} \tag{6.71}$$

Note that they must be somehow associated with the $PSL(2,\mathbb{R})$ subgroup of *Diff*$^+(S^1)$, since the $\hat{\xi}$-component of the null directions coincide with the generators of $\mathfrak{psl}(2,\mathbb{R})$.

If the $\mathring{\alpha}$-component of these null directions were equal to zero, then the reduced phase space would be simply $[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})] \times \mathring{\mathfrak{diff}}^{*}(S^1)$. Note that this is a trivial vector bundle which is locally isomorphic to the cotangent bundle of $\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$, since the cotangent space at the projection of the identity, $[I] \in \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$, is isomorphic to $\mathring{\mathfrak{diff}}^{*}(S^1)$. Based on the first approach for reducing the phase space, we actually know that the correct topology for the reduced phase space is actually the cotangent bundle of $\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$, that is, $T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$. Intuitively, we can think that the fact that the null directions are "tilted" instead of purely "horizontal" (i.e., their $\hat{\alpha}$-component is non-zero) leads to a quotient that is not just the trivial bundle $[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})] \times \mathring{\mathfrak{diff}}^{*}(S^1)$, but rather the "twisted" bundle $T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$.[27]

6.4 Removing the residual $PSL(2,\mathbb{R})$

With the insight above, we will prove that the (fully) reduced phase space is

$$\widetilde{\mathcal{P}} = T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})] \tag{6.72}$$

by showing that (i) there is a natural projection map $J : \mathit{Diff}^{+}(S^1) \times \mathring{\mathfrak{diff}}^{*}(S^1) \to T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$ and that (ii) the canonical symplectic form $\widetilde{\omega}$ on $T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$, associated with its cotangent bundle structure, which is closed and non-degenerate,

[27] It is interesting to note there is a surprising symplectomorphism between $T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$ and $[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})] \times [\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$, for the natural symplectic structures on each space, provided by the Moss map. Despite this classical equivalence, the quantization may depend on the particular "presentation" of the phase space. In our approach $T^*[\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})]$ is certainly the more natural presentation.

pulls-back to $\mathit{Diff}^{+}(S^1) \times \mathring{\mathfrak{diff}}^{*}(S^1)$, i.e., $\hat{\omega} = J^{*}\widetilde{\omega}$.

6.4.1 Proof of (i): Existence of the projection map J

We begin by proving that this projection map exists. Consider the following theorem, which we will prove below.

Theorem: Let G be a Lie group and let H be a (closed) subgroup of G. Denote the Lie algebra of G by $\mathfrak{g}$ and the Lie algebra of H by $\mathfrak{h}$ (naturally, $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$). Define the *annihilator* of $\mathfrak{h}$, denoted by $\mathring{\mathfrak{g}}^{*}$, as the subspace (not necessarily a subalgebra) of the dual Lie algebra of G, $\mathring{\mathfrak{g}}^{*} \subset \mathfrak{g}^{*}$, such that $\mathfrak{h}$ (and only $\mathfrak{h}$) is in the kernel of all elements of $\mathring{\mathfrak{g}}^{*}$. That is,

$$\mathring{\mathfrak{g}}^{*} := \{\mathring{\sigma} \in \mathfrak{g}^{*},\, \mathring{\sigma}(\xi) = 0 \text{ if } \xi \in \mathfrak{h} \subset \mathfrak{g}\} \tag{6.73}$$

Let G/H be (left) coset space of H in G and denote by $q : G \to G/H$ the corresponding quotient map, $g \mapsto [g] = gH$. Then there is a natural, well-defined quotient map $J : G \times \mathring{\mathfrak{g}}^{*} \to T^{*}(G/H)$,

$$J(g, \mathring{\sigma}) = \widetilde{\sigma} \tag{6.74}$$

where $g \in G$, $\mathring{\sigma} \in \mathring{\mathfrak{g}}^{*}$ and $\widetilde{\sigma} \in T^{*}_{[g]}(G/H)$ is defined by

$$\widetilde{\sigma}(\widetilde{X}) := \mathring{\sigma}(\Xi(X)) \tag{6.75}$$

where $\widetilde{X} \in T_{[g]}(G/H)$ is a tangent vector at $[g]$, $X \in T_g G$ is any vector at g in the pre-image of $\widetilde{X}$ under q_* (i.e., $q_* X = \widetilde{X}$), and Ξ is the Maurer-Cartan form[28]. It is interesting to visualize this map as part of the following commutative diagram

$$\begin{array}{ccc} G \times \mathring{\mathfrak{g}}^* & \xrightarrow{\ J\ } & T^*(G/H) \\ {\scriptstyle p_1}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ G & \xrightarrow{\quad q \quad} & G/H \end{array}$$

where $p_1 : G \times \mathring{\mathfrak{g}}^* \to G$ is the projector on the first Cartesian factor, $p_1(g, \sigma) = g$, and $\pi : T^*(G/H) \to G/H$ is the projection map of the cotangent bundle (which maps a dual vector to the point it is based at).

We must show that the map J is well-defined and surjective.[29] The only arbitrariness in the definition of J is the choice of X in the pre-image of $\widetilde{X}$ under q_*, so we must show that if we choose another X' satisfying $q_* X' = \widetilde{X}$ then the right-hand side of (6.75) is unchanged. First let us characterize the kernel of q_*, i.e., given a point $g \in G$, what are the tangent vectors $\gamma \in T_g G$ such that $q_* \gamma = 0$? If $t \mapsto \Gamma_t$ is a curve starting at g and tangent to γ, the condition $q_* \gamma = 0$ implies that, to first order in t, Γ_t must be projected to a fixed point, i.e., $q(\Gamma_t) \approx [g]$. This implies that there must exist a curve $t \mapsto h_t$ in H, starting at e, such that $\Gamma_t \approx g h_t$, where $\approx$ means up to first order in t (i.e., the two curves define the same tangent vector at $t = 0$). But this implies that $\gamma = l_{g*} \zeta$ for some $\zeta \in T_e H \sim \mathfrak{h}$. The conclusion is

[28]The Maurer-Cartan form, Ξ, is a $\mathfrak{g}$-valued 1-form on G defined as follows. Let $l_g : G \to G$ be the left translation by g, i.e., $l_g(g') := gg'$. Given $X \in T_g G$, we have $\Xi(X) := l_{g^{-1}*} X$, where $l_{g^{-1}*} X \in T_e G \sim \mathfrak{g}$.

[29]Rigorously, a map f is a *quotient map* if it is surjective and the topology of the codomain is induced from f (i.e., U is open iff $f^{-1}(U)$ is open). We are only proving surjection here, but subsequently we will show that $\ker(J_*)$ has constant dimension which can be used to prove this topological condition.

therefore

$$q_*X = 0 \quad \text{iff} \quad \Xi(X) \in \mathfrak{h} \subset \mathfrak{g} \tag{6.76}$$

Given a vector $\widetilde{X}$ at $[g]$, let X and X' be two vectors at g in the pre-image of $\widetilde{X}$ under q_*, i.e., $q_*X = q_*X' = \widetilde{X}$. Thus $q_*(X' - X) = 0$ and, from the result above, $\Xi(X' - X) \in \mathfrak{h}$. That is, $\Xi(X') = \Xi(X) + \zeta$ for some $\zeta \in \mathfrak{h}$. If we use X' in the right-hand side of (6.75), we get $\sigma(\Xi(X')) = \sigma(\Xi(X)) + \sigma(\zeta) = \sigma(\Xi(X))$, since $\sigma \in \mathring{\mathfrak{g}}^*$ annihilates the $\mathfrak{h}$ subspace, which yields the same result as if we had used X. We conclude that J is well-defined. To prove surjectiveness, we must show that for any $\widetilde{\sigma} \in T^*(G/H)$, there exists a point $(g, \sigma) \in G \times \mathring{\mathfrak{g}}^*$ such that $J(g, \sigma) = \widetilde{\sigma}$. Since $q : G \to G/H$ is surjective, there is always a $g \in G$ such that $\pi(\widetilde{\sigma}) = [g]$. Now let $\widetilde{\sigma}$ be a generic dual vector at $[g]$ on G/H and pull it back to g, $q^*\widetilde{\sigma}$, and further pull it back to e, $l_g^* q^* \widetilde{\sigma}$. This defines an element of $\mathring{\mathfrak{g}}^*$ because, given any $\zeta \in \mathfrak{h}$ we have that $q_* l_{g*} \zeta = 0$, so $l_g^* q^* \widetilde{\sigma}(\zeta) = \widetilde{\sigma}(q_* l_{g*} \zeta) = 0$. Thus, for any $\widetilde{X} \in T_{[g]}(G/H)$,

$$J(g, l_g^* q^* \widetilde{\sigma})(\widetilde{X}) = l_g^* q^* \widetilde{\sigma}(\Xi(X)) = \widetilde{\sigma}(q_* l_{g*} l_{g^{-1}*} X) = \widetilde{\sigma}(q_* X) = \widetilde{\sigma}(\widetilde{X}) \tag{6.77}$$

which means that $(g, l_g^* q^* \widetilde{\sigma})$ is in the pre-image of $\widetilde{\sigma}$ under J, so we conclude that J is surjective. This finishes the proof of the theorem.

An interesting property of the map J is that it has constant kernel dimension equal to $\dim(\mathfrak{h})$, or more precisely, there is an isomorphism between $\mathfrak{h}$ and $\ker(J_*)$ everywhere in $G \times \mathring{\mathfrak{g}}^*$. This follows from the fact that, in the course of proving the theorem above, we have constructed an isomorphism between $T^*_{[g]}(G/H)$ and $\mathring{\mathfrak{g}}^*$. In particular, given any $g \in G$, we can see that the map $\sigma \in \mathring{\mathfrak{g}}^* \mapsto \widetilde{\sigma} \in T^*_{[g]}(G/H)$, where $\widetilde{\sigma}(\widetilde{X}) := \sigma(X)$, for any X satisfying $q_*X = \widetilde{X}$, and the map $\widetilde{\sigma} \in T^*_{[g]}(G/H) \mapsto l^*_g q^* \widetilde{\sigma} \in \mathring{\mathfrak{g}}^*$ are actually inverses of each other. Therefore, since the fibers of these two bundles are isomorphic and the base space G is collapsed to G/H, we see that a number $\dim(\mathfrak{h})$ of directions are mapped to zero at each point. In fact, we can derive an explicit formula for $\ker(J_*)$. Let $\eta = (X, \alpha) \in T_gG \oplus \mathring{\mathfrak{g}}^*$ be a vector in $\ker(J_*)$ at $(g, \sigma) \in G \times \mathring{\mathfrak{g}}^*$,

$$J_*\eta = J_*(X, \alpha) = 0 \tag{6.78}$$

From the commutative diagram, $\pi \circ J = q \circ p_1$, we get $\pi_* J_* \eta = 0 = q_* p_{1*} \eta = q_* X$, so from (6.76) we have $\Xi(X) \in \mathfrak{h}$. Let $X = l_{g*}\zeta$, where $\zeta \in \mathfrak{h}$, and let $t \mapsto h_t$ be a curve in $H \subset G$ starting at e and tangent to ζ. Thus the curve $t \mapsto gh_t$ in G starts at g and is tangent to X. Also, the curve $t \mapsto \sigma + t\alpha$ in $\mathring{\mathfrak{g}}^*$ starts at σ and is tangent to α. Thus the curve $t \mapsto (gh_t, \sigma + t\alpha)$ in $G \times \mathring{\mathfrak{g}}^*$ starts at (g, σ) and is tangent to $\eta = (X, \alpha)$. Since this curve projects entirely to the fiber over $[g]$ under J, we can define the pushed vector by the derivative

$$J_*(X, \alpha) = \frac{d}{dt} J(gh_t, \sigma + t\alpha)\bigg|_{t=0} \tag{6.79}$$

where we are making use of the natural identification between *vertical* vectors in $T_{\widetilde{\sigma}}[T^*(G/H)]$ and the fiber itself $T^*_{\pi(\widetilde{\sigma})}(G/H)$. Let $\widetilde{Y}$ be a vector at $[g]$ and let Y

is a vector at g satisfying $q_* Y = \widetilde{Y}$. We note that the vector $Y_t := r_{h_t *} Y$, where r_g is the right-translation in G by g, is a vector at gh_t satisfying $q_* Y_t = \widetilde{Y}$. This follows from the fact that, as $[g] = [gh]$ for $h \in H$, then $q = q \circ r_h$, which implies that $q_* Y_t = q_* r_{h_t *} Y = q_* Y = \widetilde{Y}$. Consequently,

$$J(gh_t, \sigma + t\alpha)(\widetilde{Y}) = (\sigma + t\alpha)\left(\Xi(Y_t)\right) = (\sigma + t\alpha)\left(l_{(gh_t)^{-1}*} r_{h_t *} Y\right) \tag{6.80}$$

Since the right and left translations commute, the argument of the dual vector above can be written as $l_{(gh_t)^{-1}*} r_{h_t *} Y = l_{h_t^{-1}*} l_{g^{-1}*} r_{h_t *} Y = l_{h_t^{-1}*} r_{h_t *} l_{g^{-1}*} Y$.

Before further manipulations of the expression above, let us introduce the *adjoint map*, a recurrent object for the remainder of this paper. Given a group G the *adjoint action* of $g \in G$ on $g' \in G$ is defined as

$$\mathrm{Ad}_g(g') = g g' g^{-1} \tag{6.81}$$

Noting that $\mathrm{Ad}_g(e) = e$, the push-forward operation maps $T_e G$ into itself. This gives rise to a natural action of G on its Lie algebra $\mathfrak{g}$, which we also refer to as the adjoint action of G on $\mathfrak{g}$ and denote as

$$\mathrm{ad}_g \xi := \mathrm{Ad}_{g*} \xi \tag{6.82}$$

where $\xi \in \mathfrak{g} \sim T_e G$. For a fixed element ξ, we can think of $\mathrm{ad}_g \xi$ as a map from G

into $\mathfrak{g}$, that is, we can define $\mathrm{ad}\,\xi(g) := \mathrm{ad}_g\xi$. The push-forwards of this map sends a vector $\eta \in T_eG \sim \mathfrak{g}$ to a vector in $T_\xi\mathfrak{g}$; but since $\mathfrak{g}$ is a vector space, any of its tangent spaces can be naturally identified with $\mathfrak{g}$ itself; therefore $(\mathrm{ad}\,\xi)_*$ can be seen as a map from $\mathfrak{g}$ to itself. We then define the adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$ as

$$\mathrm{ad}_\eta\xi := (\mathrm{ad}\,\xi)_*\eta \tag{6.83}$$

This adjoint action is directly related to the Lie algebra product,

$$\mathrm{ad}_\eta\xi = [\eta, \xi] \tag{6.84}$$

While it is straightforward to show this using the formal definition of the Lie product in terms of the Lie brackets of left-invariant vector fields, we can trivially see it from a matrix realization of G: in this case the group exponential coincides with the matrix exponential, $\exp(\eta) = e^\eta$, and a vector tangent to a matrix-valued curve is simply given by the standard parameter-derivative of this curve; then $\mathrm{ad}_\eta\xi = (\mathrm{ad}\,\xi)_*\eta = \frac{d}{dt}\mathrm{ad}\,\xi(e^{t\eta}) = \frac{d}{dt}\mathrm{ad}_{e^{t\eta}}\xi = \frac{d}{dt}e^{t\eta}\xi e^{-t\eta} = \eta\xi - \xi\eta = [\eta, \xi]$, where the t-derivatives are all evaluated at $t = 0$.

Let us return to the expression $l_{h_t^{-1}*}r_{h_t*}l_{g^{-1}*}Y$. Note that $l_g \circ r_{g^{-1}} = r_{g^{-1}} \circ l_g = \mathrm{Ad}_g$. Therefore $l_{h_t^{-1}*}r_{h_t*}l_{g^{-1}*}Y = \mathrm{Ad}_{h_t^{-1}*}\Xi(Y) = \mathrm{ad}_{h_t^{-1}}\Xi(Y)$, which gives

$$J(gh_t, \sigma + t\alpha)(\widetilde{Y}) = (\sigma + t\alpha)\left(\mathrm{Ad}_{h_t^{-1}*}\Xi(Y)\right) \tag{6.85}$$

so

$$\left(J_*(X,\alpha)\right)(\widetilde{Y}) = \alpha(\Xi(Y)) + \frac{d}{dt}\sigma\left(\mathrm{Ad}_{h_t^{-1}*}\Xi(Y)\right)\bigg|_{t=0} = \alpha(\Xi(Y)) - \sigma\left(\mathrm{ad}_\zeta\Xi(Y)\right) \tag{6.86}$$

where we have used that h_t^{-1} is tangent to $-\zeta$ at $t = 0$, so $\frac{d}{dt}\mathrm{Ad}_{h_t^{-1}*}\Xi(Y) = \frac{d}{dt}\mathrm{ad}_{h_t^{-1}}\Xi(Y) = (\mathrm{ad}\,\Xi(Y))_*(-\zeta) = -\mathrm{ad}_\zeta\Xi(Y)$, where $\frac{d}{dt}$ is evaluated at $t = 0$.

We will also define the *coadjoint action* of $g \in G$ on $\sigma \in \mathfrak{g}^*$ by

$$\mathrm{coad}_g\sigma := \mathrm{Ad}^*_{g^{-1}}\sigma \tag{6.87}$$

where on the right-hand side σ is seen as a 1-form at e. The reason for taking the inverse of g in Ad^* is so that the coadjoint action composes nicely, i.e., $\mathrm{coad}_g \circ \mathrm{coad}_{g'} = \mathrm{coad}_{gg'}$. Also, analogous to the definition of the adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$, we define the coadjoint action of $\mathfrak{g}$ on $\mathfrak{g}^*$ by

$$\mathrm{coad}_\zeta\sigma := \frac{d}{dt}\mathrm{coad}_{g_t}\sigma\bigg|_{t=0} \tag{6.88}$$

where $t \mapsto g_t$ is a curve in G starting at e and tangent to $\zeta \in \mathfrak{g}$. It is worth noting, for later reference, that $\mathrm{coad}_g\sigma(\xi) = \mathrm{Ad}^*_{g^{-1}}\sigma(\xi) = \sigma(\mathrm{Ad}_{g^{-1}*}\xi) = \sigma(\mathrm{ad}_{g^{-1}}\xi)$ and, by taking g as a curve g_t tangent to a vector η at e and evaluating the t-derivative we get that

$$\mathrm{coad}_\eta\sigma(\xi) = -\sigma(\mathrm{ad}_\eta\xi) \tag{6.89}$$

revealing that coad_η is *minus* the algebraic dual of ad_η.

Returning again to the evaluation of J_*, we can equivalently write (6.86) as

$$(J_*(X,\alpha))\,(\widetilde{Y}) = \big(\alpha + \mathrm{coad}_{\Xi(X)}\sigma\big)\,(\Xi(Y)) \tag{6.90}$$

recalling that $\Xi(X) = \zeta$ is the vector tangent to $t \mapsto h_t$. So, given the identification $\mathrm{Ver}(T_{\widetilde{\sigma}}[T^*(G/H)]) \sim T^*_{\pi(\widetilde{\sigma})}(G/H)$, we have

$$J_*(l_{g*}\zeta, \alpha) \cong J(g, \alpha + \mathrm{coad}_\zeta \sigma) \tag{6.91}$$

where $\zeta \in \mathfrak{h}$. Therefore, $\eta = (X,\alpha) \in \ker(J_*)$ if and only if $\Xi(X) \in \mathfrak{h}$ and

$$\alpha = -\mathrm{coad}_{\Xi(X)}\sigma \tag{6.92}$$

This confirms that, at every point $(g,\sigma) \in G \times \mathring{\mathfrak{g}}^*$, the kernel of J_* is isomorphic to $\mathfrak{h}$ via the map

$$\mathfrak{h} \to \ker(J_*)\,, \quad \zeta \mapsto \eta = (l_{g*}\zeta, -\mathrm{coad}_\zeta \sigma) \tag{6.93}$$

Note that these collapsing directions are "tilted" in the bundle $G \times \mathring{\mathfrak{g}}^*$, in the sense that $p_{2*}\eta \neq 0$, which is reminiscent of the the null directions in (6.71).

We are interested in the particularization of the theorem for $G = \mathit{Diff}^+(S^1)$ and $H = \mathit{PSL}(2,\mathbb{R})$. Let us express it in the notation we have been using before. The theorem states that there exists a projection map $J : \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1) \to$

$T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$ defined by

$$J(\psi, \mathring{\sigma}) = \widetilde{\sigma}\,, \quad \widetilde{\sigma}(\widetilde{\eta}) = \mathring{\sigma}(l_{\psi^{-1}*}\hat{\eta}) \tag{6.94}$$

where $\widetilde{\eta} \in T_{[\psi]}(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$ and $\hat{\eta} \in T_\psi \mathit{Diff}^+(S^1)$ is any vector satisfying $q_*\hat{\eta} = \widetilde{\eta}$ (with $q : \mathit{Diff}^+(S^1) \to \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ being the quotient map). The kernel of J_* is given by

$$\hat{\eta} \in \ker(J_*) \quad \text{iff} \quad \hat{\eta} = (\zeta, -\mathrm{coad}_\zeta \mathring{\sigma})\,,\ \zeta \in \mathfrak{psl}(2,\mathbb{R}) \tag{6.95}$$

Note a slight difference in notation: as explained in footnote 26, a tangent vector at a point $(\psi, \mathring{\sigma}) \in \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$ is being characterized as $(\xi, \hat{\alpha}) \in \mathfrak{diff}(S^1) \oplus \mathring{\mathfrak{diff}}^*(S^1)$; but during the discussion of the above theorem a tangent vector at $(g,\sigma) \in G \times \mathring{\mathfrak{g}}^*$ was characterized as $(X, \alpha) \in T_gG \oplus \mathring{\mathfrak{g}}^*$. Of course, these two characterizations are naturally equivalent since $\mathfrak{g}$ can be identified with T_gG via the map l_{g*}; moreover, despite the use of similar notation, we expect that the context should prevent any confusion as it should be clear whether the first component of the pair belongs to $\mathfrak{g}$ or T_gG. We can explicitly compute the coadjoint action using the relation (6.89),

$$\mathrm{coad}_\zeta \mathring{\sigma}(\xi) = -\mathring{\sigma}(\mathrm{ad}_\zeta \xi) = -\mathring{\sigma}([\zeta, \xi]) \tag{6.96}$$

where $\zeta \in \mathfrak{psl}(2,\mathbb{R})$ and $\xi \in \mathfrak{diff}(S^1)$. As discussed previously, we can characterize

$\mathfrak{diff}(S^1)$ and $\mathfrak{diff}^*(S^1)$ using vector fields and quadratic form fields on S^1, so we write $\mathring{\sigma} = \sigma d\theta^2$, $\zeta = z\partial_\theta$ and $\xi = f\partial_\theta$. Before we proceed, we must explain an unfortunate (and potentially confusing) aspect of this notation: the Lie brackets on $\mathfrak{diff}(S^1)$, $[\xi, \xi']_{\mathfrak{diff}}$, differs by a sign from the vector field brackets on S^1, $[\xi, \xi']_{S^1}$. That is,

$$[\xi, \xi']_{\mathfrak{diff}} = -[\xi, \xi']_{S^1} \tag{6.97}$$

In order to clarify this point, let us use a more explicit notation. Let $G = \mathrm{Diff}(\mathcal{M})$ be the group of diffeomorphisms on a manifold $\mathcal{M}$. Let $f_x : G \to \mathcal{M}$ be the action of $\psi \in G$ on the point $x \in \mathcal{M}$, i.e., $f_x(\psi) := \psi(x)$. As with any group action, to any element of the algebra $\xi \in \mathfrak{g}$ we can associate a vector field on $\mathcal{M}$ defined by $V_\xi|_x := f_{x*}\xi$. The group of diffeomorphisms is special in the sense that this map is an isomorphism, allowing us to identify $\mathfrak{g} \sim \mathrm{Vect}(\mathcal{M})$.[30] Since G acts on the left of $\mathcal{M}$, there is an anti-homomorphism between the Lie algebra of G and the vector field algebra of $\mathcal{M}$, that is, $[V_\xi, V_{\xi'}] = V_{[\xi',\xi]}$. This is perhaps aesthetically unpleasing, but completely transparent. Confusion may arise, however, when we start writing ξ as a shorthand for V_ξ. In this case, one solution is to always specify which bracket is being used by writing $[\xi, \xi']_{\mathfrak{diff}} := [\xi, \xi']$ and $[\xi, \xi']_{\mathcal{M}} := [V_\xi, V_{\xi'}]$, so we would have

[30]The are some subtleties associated with the infinite dimensionality of the group of diffeomorphisms, such as the fact that the exponential map is not surjective in any neighborhood of the identity (although it has dense image). So we can think of this identification of $\mathfrak{diff}(\mathcal{M})$ with $\mathrm{Vect}(\mathcal{M})$ as a practical/formal characterization of $\mathfrak{diff}(\mathcal{M})$.

$[\xi, \xi']_{\mathfrak{diff}} = -[\xi, \xi']_{\mathcal{M}}$. In this manner, we can write (6.96) more precisely as

$$\mathrm{coad}_\zeta \mathring{\sigma}(\xi) = -\mathring{\sigma}(\mathrm{ad}_\zeta \xi) = -\mathring{\sigma}([\zeta, \xi]_{\mathfrak{diff}}) = \mathring{\sigma}([\zeta, \xi]_{S^1}) \tag{6.98}$$

Thus, if we denote $\mathrm{coad}_\zeta \mathring{\sigma} = \sigma' d\theta^2$, we have

$$\int d\theta\, \sigma' f = \int d\theta\, \sigma \left(z\frac{\partial f}{\partial \theta} - f\frac{\partial z}{\partial \theta} \right) = \int d\theta \left(-z\frac{\partial \sigma}{\partial \theta} - 2\sigma\frac{\partial z}{\partial \theta} \right) f \tag{6.99}$$

Since this is valid for any $\xi \in \mathfrak{diff}(S^1)$ (i.e., for any $f \in C^\infty(S^1, \mathbb{R})$), we conclude that

$$\mathrm{coad}_\zeta \mathring{\sigma} = -\left(z\frac{\partial \sigma}{\partial \theta} + 2\sigma\frac{\partial z}{\partial \theta} \right) d\theta^2 \tag{6.100}$$

Therefore we see that the collapsed directions under the projection J (i.e., the kernel of J_*) are given, according to (6.95), by

$$\hat{\eta} = \big(z\partial_\theta, (z\partial_\theta \sigma + 2\sigma\partial_\theta z) d\theta^2 \big) \tag{6.101}$$

where z is any linear combination of 1, $\sin\theta$, $\cos\theta$. These directions precisely match the null directions (6.71) of $\hat{\omega}$ on $\hat{\mathcal{S}}$, confirming that J is the desired quotient map to the reduced phase space.

6.4.2 Proof of (ii): The natural symplectic form is the physical one

Now that we have shown that the reduced phase space $\widetilde{\mathcal{P}}$ has topology $T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$, we must figure out what is the correct symplectic structure on it. The symplectic form $\widetilde{\omega}$ on $\widetilde{\mathcal{P}}$ must be such that $J^*\widetilde{w} = \hat{\omega}$, where $\hat{\omega}$ is the symplectic form on $\hat{\mathcal{S}}$. The natural guess is the canonical symplectic form associated with the cotangent bundle structure of $T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$. Let us review how it is constructed and then show that this is indeed the correct choice.

Let $T^*\mathcal{M}$ be the cotangent bundle of a manifold $\mathcal{M}$ and let $\pi : T^*\mathcal{M} \to \mathcal{M}$ be the projection map (which maps a dual vector to the point it is based at). There is a natural 1-form on $T^*\mathcal{M}$, called *canonical potential 1-form* θ, defined as

$$\theta(\eta) := \sigma(\pi_*\eta) \tag{6.102}$$

where $\eta \in T_\sigma(T^*\mathcal{M})$. In words, "given a vector η at $\sigma \in T^*\mathcal{M}$, we project it down to $\mathcal{M}$ and act with σ on it". Taking the exterior derivative of θ yields a 2-form,

$$\omega = d\theta \tag{6.103}$$

which is called the *canonical symplectic 2-form* on $T^*\mathcal{M}$. Note that it is indeed a symplectic form since it is closed and non-degenerate.

So let θ be canonical potential 1-form on $T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$ and let us evaluate its pull-back to $\hat{\mathcal{S}}$, $\hat{\theta} = J^*\theta$. If $\hat{\eta}$ is a vector at $(\psi, \mathring{\sigma}) \in \hat{\mathcal{S}}$, then

$$\hat{\theta}(\hat{\eta}) = \theta(J_*\hat{\eta}) = J(\psi, \mathring{\sigma})(\pi_* J_* \hat{\eta}) = J(\psi, \mathring{\sigma})(q_* p_{1*} \hat{\eta}) \tag{6.104}$$

where we have used the definition of θ and the relation $\pi \circ J = q \circ p_1$. But note that $p_{1*}\hat{\eta}$ is obviously a vector at $\psi \in \mathit{Diff}^+(S^1)$ which projects to $q_* p_{1*}\hat{\eta}$ under q_*, so by the definition of J we have

$$\hat{\theta}(\hat{\eta}) = \mathring{\sigma}(l_{\psi^{-1}*} p_{1*} \hat{\eta}) \tag{6.105}$$

If $\hat{\eta}$ is expressed as $(\xi, \hat{\alpha}) \in \mathfrak{diff}(S^1) \oplus \mathring{\mathfrak{diff}}^*(S^1)$, we have that $p_{1*}\hat{\eta} = l_{\psi *}\xi$, so

$$\hat{\theta}(\hat{\eta}) = \mathring{\sigma}(\xi) \tag{6.106}$$

Since the exterior derivative commutes with the pull-back, we have $\hat{\omega} := J^*(\omega) = J^*(d\theta) = d(J^*\theta) = d\hat{\theta}$. The exterior derivative of a 1-form is generally given by

$$d\hat{\theta}(\hat{\eta}, \hat{\eta}') = \hat{\eta}[\hat{\theta}(\hat{\eta}')] - \hat{\eta}'[\hat{\theta}(\hat{\eta})] - \hat{\theta}([\hat{\eta}, \hat{\eta}']) \tag{6.107}$$

where $\hat{\eta}[\hat{\theta}(\hat{\eta}')]$ denotes the directional derivative of the scalar field $\hat{\theta}(\hat{\eta}')$ along the vector $\hat{\eta}$, and $[\hat{\eta}, \hat{\eta}']$ denotes the Lie brackets of the vector fields $\hat{\eta}$ and $\hat{\eta}'$. Note that since $d\hat{\theta}$ is a tensor it depends only on the values of $\hat{\eta}$ and $\hat{\eta}'$ at the point where it is evaluated, so the right-hand side of this formula is independent of how $\hat{\eta}$ and $\hat{\eta}'$ are extended to vector fields in a neighborhood of the evaluation point. A particularly convenient choice is to take $\hat{\eta} = (\xi, \hat{\alpha})$ and $\hat{\eta}' = (\xi', \hat{\alpha}')$ with ξ, ξ', $\hat{\alpha}$ and $\hat{\alpha}$ being constants in a neighborhood of $(\psi, \mathring{\sigma})$. In other words, $\hat{\eta}$ is extended into a vector field in such a way that $p_{1*}\hat{\eta}$ is a left-invariant field on $\mathit{Diff}^+(S^1)$ and $p_{2*}\hat{\eta}$ is a fixed point in $\mathring{\mathfrak{diff}}^*(S^1)$. With this choice we note that $\hat{\theta}(\hat{\eta}') = \mathring{\sigma}(\xi')$ is a function on $\mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$ depending only on $\mathring{\sigma}$ (not on ψ). Thus,

$$
\begin{aligned}
\hat{\eta}[\hat{\theta}(\hat{\eta}')] &= (\xi, \hat{\alpha})[\mathring{\sigma}(\xi')] \\
&= (\xi, 0)[\mathring{\sigma}(\xi')] + (0, \hat{\alpha})[\mathring{\sigma}(\xi')] \\
&= \hat{\alpha}(\xi') \qquad (6.108)
\end{aligned}
$$

where the first term on the second line vanishes because we can take a horizontal curve $t \mapsto (\psi_t, \mathring{\sigma})$ as tangent to $(\xi, 0)$, so that $\frac{d}{dt}\mathring{\sigma}(\xi') = 0$; for the second term on the second line we can take the vertical curve $t \mapsto (\psi, \mathring{\sigma} + t\hat{\alpha})$ as tangent to $(0, \hat{\alpha})$, so that $\frac{d}{dt}(\mathring{\sigma} + t\hat{\alpha})(\xi') = \hat{\alpha}(\xi')$. Analogously, $\hat{\eta}'[\hat{\theta}(\hat{\eta})] = \hat{\alpha}'(\xi)$. Lastly, consider the term $\hat{\theta}([\hat{\eta}, \hat{\eta}'])$ which is given by $\mathring{\sigma}(l_{\psi^{-1}*}p_{1*}[\hat{\eta}, \hat{\eta}'])$. Given our choice, we have $p_{1*}\hat{\eta} = l_{\psi*}\xi$ and $p_{1*}\hat{\eta}' = l_{\psi*}\xi'$, so that both fields are projected nicely[31] to $\mathit{Diff}^+(S^1)$, which implies that $p_{1*}[\hat{\eta}, \hat{\eta}'] = [p_{1*}\hat{\eta}, p_{1*}\hat{\eta}'] = [l_{\psi*}\xi, l_{\psi*}\xi'] = l_{\psi*}[\xi, \xi']_{\mathfrak{diff}}$, where in the last step we used the definition of the Lie algebra product as being homomorphic to the

[31]Note that, in general, vector fields cannot be pushed through a (non-injective) map. Let $f : \mathcal{M} \to \mathcal{N}$ be a smooth map and let V be a vector field on $\mathcal{M}$. For each $x \in \mathcal{M}$ we can define the push-forward f_*V_x to $y = f(x) \in \mathcal{N}$. But if f is not injective, there may exist $x' \neq x$ such that $f(x') = y$, and in general the push-forward $f_*V_{x'}$ to y will not coincide with f_*V_x. If a vector field V on $\mathcal{M}$ satisfies $f_*V_x = f_*V_{x'}$ for all x, $x' \in f^{-1}(y)$, for all $y \in \mathcal{N}$, then we say that V is projected *nicely* by f to $\mathcal{N}$. The Lie brackets behave nicely under nice projections, in the sense that $f_*[V, V'] = [f_*V, f_*V']$.

Lie bracket algebra of left-invariant fields. Thus,

$$\hat{\theta}([\hat{\eta}, \hat{\eta}']) = \mathring{\sigma}([\xi, \xi']_{\mathfrak{diff}}) = -\mathring{\sigma}([\xi, \xi']_{S^1}) \tag{6.109}$$

So we have

$$\hat{\omega}(\hat{\eta}, \hat{\eta}') = \hat{\alpha}(\xi') - \hat{\alpha}'(\xi) + \mathring{\sigma}([\xi, \xi']_{S^1}) \tag{6.110}$$

which matches exactly with the symplectic form on $\hat{\mathcal{S}}$ given in (6.66), confirming that the correct symplectic form $\widetilde{\omega}$ on $\widetilde{\mathcal{P}}$ is precisely the canonical symplectic form

$$\widetilde{\omega} = \omega = d\theta \tag{6.111}$$

where θ is the canonical potential 1-form on $T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$. This concludes the reduction process from $\mathcal{P} = \mathrm{Riem}(\Sigma, \gamma) \times \mathrm{Sym}(\Sigma, (2, 0))$ to $\widetilde{\mathcal{P}} = T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$.

6.5 Diagrammatic summary

In Fig. 4 we illustrate the reduction process via conformal coordinates in a diagrammatic way.

7 The reduced Hamiltonian

The Hamiltonian description of a dynamical system involves the specification of a phase space (its topology and symplectic structure) and a Hamiltonian function which generates time evolution. Having already completed the specification of the

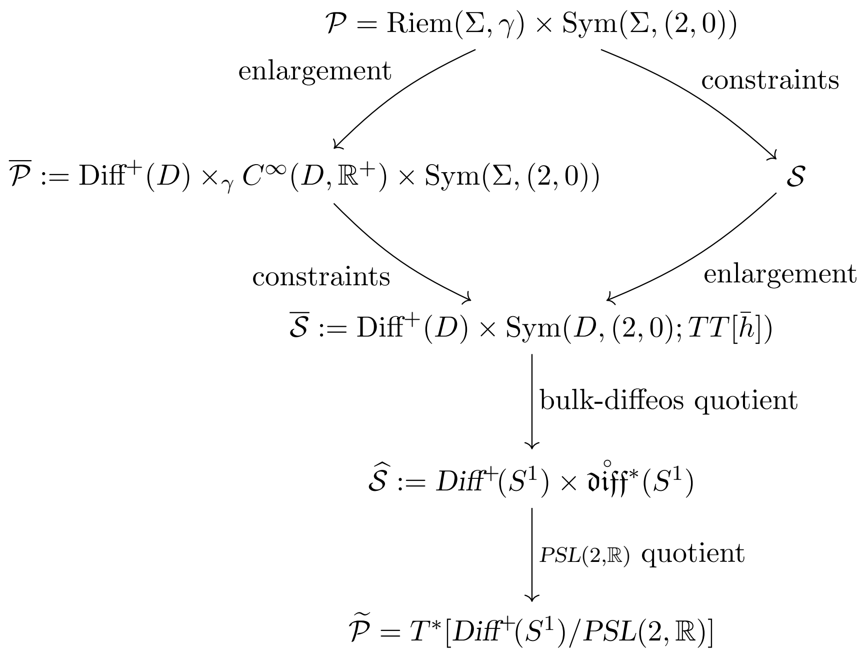


Figure 4: A diagrammatic summary of the phase space reduction process for the diamond. From $\mathcal{P}$ to $\bar{\mathcal{P}}$ the phase space is enlarged by introducing "conformal coordinates" $(\Psi, \Omega, \bar{\sigma}^{ab})$, and the constraints define a submanifold $\bar{\mathcal{S}}$. The order of these two steps can be interchanged, first going to $\mathcal{S}$ by imposing the constraints and then enlarging to $\bar{\mathcal{S}}$ by introducing conformal coordinates $(\Psi, \bar{\sigma}^{ab})$. From $\bar{\mathcal{S}}$ to $\hat{\mathcal{S}}$ the bulk diffeomorphisms are removed, and to $\widetilde{\mathcal{P}}$ the remaining $PSL(2,\mathbb{R})$ action is quotiented out.

reduced phase space $\widetilde{\mathcal{P}}$ for the diamond, we now proceed to computing the Hamiltonian H which generates time-evolution according to our gauge-fixed choice of time, i.e., evolution along the CMC foliation.

7.1 General review

First, let us recall how every symplectic motion on a phase space corresponds to an action principle. Let $\mathcal{P}$ be a phase space with symplectic 2-form ω. Locally, let θ be the symplectic potential 1-form for ω, i.e., $\omega = d\theta$. Given any function $H : \mathcal{P} \to \mathbb{R}$,

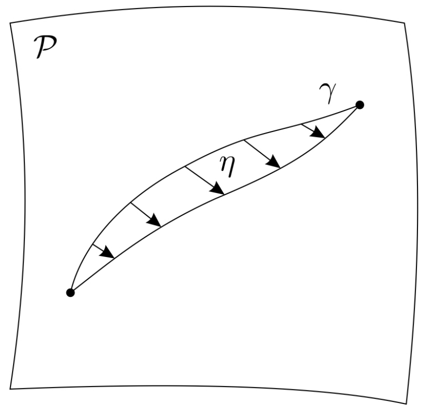


Figure 5: A curve γ on the phase space $\mathcal{P}$ is deformed by the vector field η, while its endpoints remain fixed.

define the action principle

$$S[\gamma] = \int_\gamma (\theta - H dt) \tag{7.1}$$

where $\gamma : [0,1] \to \mathcal{P}$ is a curve on the phase space.[32] The domain of S is given by all smooth curves γ with fixed endpoints. Note that this is just a covariant way to write the usual phase space action principle: from Darboux's theorem there are always local coordinates $\{p_i, q^i\}$ such that $\omega = dp_i \wedge dq^i$ and $\theta = p_i dq^i$; then if $\gamma(t) = (p_i(t), q^i(t))$, so the tangent vector is $\dot{\gamma} = \dot{p}_i \frac{\partial}{\partial p_i} + \dot{q}^i \frac{\partial}{\partial q^i}$, we have $S[\gamma] = \int_0^1 (p_i \dot{q}^i - H) dt = \int_0^1 L dt$, where L is the Lagrangian associated with H. Now we consider an infinitesimal deformation of γ by a vector field η and evaluate the corresponding variation of the action. [See Fig. 5.] More precisely, if $\Phi_\eta(s)$ is the

[32]Rigorously, in order to think of dt as a legitimate 1-form, the integral should be defined on the extended phase space, $\mathcal{P} \times \mathbb{R}$, where the "time" is adjoined to the usual phase space. In this way, the integral is evaluated along the lifted curve $\Gamma(t) := (\gamma(t), t)$. However we shall not be so pedantic about this.

diffeomorphism associated with flowing along η by a parameter-length s, define

$$\delta S := \frac{d}{ds} S[\Phi_\eta(s) \circ \gamma] \bigg|_{s=0} \tag{7.2}$$

From the definition of S we have

$$\frac{d}{ds} S[\Phi_\eta(s) \circ \gamma] = \frac{d}{ds} \int_{\Phi_\eta(s) \circ \gamma} (\theta - H dt) = \frac{d}{ds} \int_\gamma \Phi_\eta(s)^* \, (\theta - H dt) \tag{7.3}$$

and since the derivative is evaluated at $s = 0$ this gives a Lie derivative,

$$\delta S = \int_\gamma \pounds_\eta \, (\theta - H dt) \tag{7.4}$$

We have

$$\begin{aligned} \pounds_\eta \theta &= \imath_\eta d\theta + d\imath_\eta \theta \\ \pounds_\eta (H dt) &= (\pounds_\eta H) dt = (\imath_\eta dH) dt \end{aligned} \tag{7.5}$$

which leads to

$$\delta S = \int_\gamma (\imath_\eta \omega - (\imath_\eta dH) dt) + \imath_\eta \theta \big|_{\partial \gamma} \tag{7.6}$$

where the boundary term vanishes because $\eta = 0$ at the endpoints of γ (recall that we are only allowing deformations that keep the endpoints fixed). This integral can be expressed as

$$\delta S = \int_0^1 dt \, (\imath_{\dot\gamma} \imath_\eta \omega - \imath_\eta dH) \tag{7.7}$$

where $\dot\gamma$ is the vector field tangent to γ. Thus we have

$$\delta S = -\int_0^1 dt \, \imath_\eta \, (\imath_{\dot\gamma}\omega + dH) \tag{7.8}$$

The variational principle says that δS must vanish for all first-order deformations of the curve γ, i.e., for all η, so we conclude that

$$dH = -\imath_{\dot\gamma}\omega \tag{7.9}$$

That is, the time evolution associated with the action S (i.e., the curve $\gamma(t)$ satisfying the equations of motion) is precisely the symplectic motion generated by the Hamiltonian H.

7.2 The CMC Hamiltonian

We have learned that if one starts with an action S which can be written in the form (7.1) for some phase space function H, then H is the Hamiltonian which generates t-evolution. This tells us how to obtain the Hamiltonian on the reduced phase space. We begin with the Einstein-Hilbert action (3.10), written in the ADM form,

$$S = \int dt \left(\int_\Sigma d^2x \, \pi^{ab} \dot h_{ab} - H \right) \tag{7.10}$$

where H is the Hamiltonian given in (3.13). The ADM decomposition must be associated with our gauge-fixed choice of foliation, by CMC slices, so t must be some monotonic function of the CMC time $\tau = -K$. However we will leave this relationship unspecified for now, simply keeping in mind that τ is some function of t, i.e., $\tau(t)$. As we must consider only paths restricted to the constraint surface, where $H = 0$, we have

$$S = \int dt \int_{\Sigma} d^2x \, \pi^{ab} \dot{h}_{ab} \tag{7.11}$$

The goal is to write this action as an integral over curves on the reduced phase space. To do so, let us first "pull-back" this action to the enlarged space $\bar{\mathcal{S}}$. Given a curve $\gamma(t) = (h_{ab}(t), \sigma^{ab}(t))$ on the contraint surface $\mathcal{S}$, there are curves $\bar{\gamma}(t) = (\Psi(t), \bar{\sigma}^{ab}(t))$ on $\bar{\mathcal{S}}$ which map into γ under the transformation (6.21), i.e., $\gamma(t) = T_t(\bar{\gamma}(t))$. Note that we have a subscript on T_t to emphasize that this map is time-dependent since T depends on τ through λ, solution of (6.19). That there are many $\bar{\gamma}$ associated to each γ is just a consequence of the fact that $\bar{\mathcal{S}}$ has more gauge ambiguities than $\mathcal{S}$. From (6.28) and (6.42) we have

$$d^2x \, \pi^{ab} = \Psi_* \left[\vartheta_{\bar{h}} \left(e^{-\lambda} \bar{\sigma}^{ab} + \frac{1}{2} \tau \bar{h}^{ab} \right) \right] \tag{7.12}$$

The "velocity" $\dot{h}_{ab}$ can be written as $\delta h_{ab}(\dot{\gamma})$, where $\dot{\gamma}$ is the vector tangent to γ. If $\dot{\bar{\gamma}} = (\xi, \alpha^{ab})$ is the vector tangent to $\bar{\gamma}$, then from a computation similar to what led to (6.36) we obtain

$$\delta h_{ab}(\dot{\gamma}) = \Psi_* \left[-\pounds_\xi (e^\lambda \bar{h}_{ab}) + \dot{\lambda} e^\lambda \bar{h}_{ab} \right] \tag{7.13}$$

where $\dot{\lambda} := \frac{d}{dt}\lambda_t$, with λ_t being the solution of (6.19) for the curve $(\Psi(t), \bar{\sigma}^{ab}(t); \tau(t))$. Putting these together we get

$$S = \int dt \int \vartheta_{\bar{h}} \left(e^{-\lambda}\bar{\sigma}^{ab} + \frac{1}{2}\tau \bar{h}^{ab} \right) \left(-\pounds_\xi (e^\lambda \bar{h}_{ab}) + \dot{\lambda} e^\lambda \bar{h}_{ab} \right) \tag{7.14}$$

where the Ψ_* goes away because Ψ is an automorphism of D and the integral is thus invariant. Distributing the product we have

$$S = \int dt \int \vartheta_{\bar{h}} \left(-\bar{\sigma}^{ab} \pounds_\xi \bar{h}_{ab} - \frac{1}{2}\tau \bar{h}^{ab} \pounds_\xi (e^\lambda \bar{h}_{ab}) + \tau \frac{d}{dt} e^\lambda \right) \tag{7.15}$$

Let us analyze each of these three terms. The first term can be written as a total spatial derivative, since $\bar{\sigma}^{ab}$ is transverse with respect to $\bar{h}_{ab}$, and so we can apply Gauss' theorem,

$$-\int \vartheta_{\bar{h}} \bar{\sigma}^{ab} \pounds_\xi \bar{h}_{ab} = -2\int \vartheta_{\bar{h}} \bar{\sigma}^{ab} \nabla_a \xi_b = -2 \int \vartheta_{\bar{h}} \nabla_a \left(\bar{\sigma}^{ab} \xi_b \right) = -2 \int d\theta \, n_a \bar{\sigma}^{ab} \xi_b \tag{7.16}$$

From the association $\mathring{\sigma}^{ab} \leftrightarrow \mathring{\sigma}$, defined in (6.60), we identify

$$-\int \vartheta_{\bar{h}} \bar{\sigma}^{ab} \pounds_\xi \bar{h}_{ab} = \mathring{\sigma}(\hat{\xi}) \tag{7.17}$$

where $\hat{\xi}$ is the restriction of ξ to ∂D. This reveals that this term depends only on the curve $\hat{\gamma}(t) := (\psi(t), \mathring{\sigma}(t))$ projected from $\bar{\mathcal{S}}$ to $\hat{\mathcal{S}}$, where $\psi(t) = \Psi(t)|_{\partial D}$ and $\bar{\sigma}^{ab}(t) \leftrightarrow \mathring{\sigma}(t)$. In particular, note that the tangent vector to $\hat{\gamma}$ is given by $\dot{\hat{\gamma}} = (\hat{\xi}, \hat{\alpha})$, where $\alpha^{ab} \leftrightarrow \hat{\alpha}$. In fact, this term really looks like a $p\dot{q}$ term on $\hat{\mathcal{S}}$. Our goal, however, is to have a $p\dot{q}$ term on $\widetilde{\mathcal{P}}$. From formula (6.106), relating $\hat{\theta} = J^*\theta$ and $\mathring{\sigma}$, we can express this term as

$$\mathring{\sigma}(\hat{\xi}) = J^*\theta(\dot{\hat{\gamma}}) = \theta(J_*\dot{\hat{\gamma}}) \tag{7.18}$$

This further reveals that this term depends only on the curve $\widetilde{\gamma} := J \circ \bar{\gamma}$ projected from $\hat{\mathcal{S}}$ to $\widetilde{\mathcal{P}}$, since $J_*\dot{\hat{\gamma}} = \dot{\widetilde{\gamma}}$ is precisely the vector tangent to $\widetilde{\gamma}$. Thus we have

$$-\int \vartheta_{\bar{h}} \bar{\sigma}^{ab} \pounds_\xi \bar{h}_{ab} = \theta(\dot{\widetilde{\gamma}}) \tag{7.19}$$

which is the desired $p\dot{q}$ term on the reduced phase space. The second term on (7.15) can be worked out as follows,

$$-\frac{1}{2}\tau \bar{h}^{ab} \pounds_\xi (e^\lambda \bar{h}_{ab}) = -\frac{1}{2}\tau \bar{h}^{ab} \left(\xi^c \nabla_c (e^\lambda) \bar{h}_{ab} + e^\lambda 2 \nabla_a \xi_b\right) = -\tau \nabla_a \left(e^\lambda \xi^a\right) \tag{7.20}$$

so again we have a total spatial derivative and we can apply Gauss' theorem,

$$-\frac{1}{2}\tau \int \vartheta_{\bar{h}} \bar{h}^{ab} \pounds_\xi (e^\lambda \bar{h}_{ab}) = -\tau \int \vartheta_{\bar{h}} \nabla_a \left(e^\lambda \xi^a\right) = -\tau \int d\theta \, n_a e^\lambda \xi^a = 0 \tag{7.21}$$

and we conclude that it vanishes since ξ is tangent to the boundary. Finally, the third term on (7.15) can be integrated by parts (in time) to yield

$$\int dt\, \tau \frac{d}{dt} \int \vartheta_{\bar{h}} e^{\lambda} = -\int dt\, \frac{d\tau}{dt} \int \vartheta_{\bar{h}} e^{\lambda} + \left[\tau \int \vartheta_{\bar{h}} e^{\lambda} \right]_0^1 \tag{7.22}$$

Therefore, up to a time-boundary term (which does not affect the equations of motion), we have cast the action in the form

$$S[\gamma] = \int_{\widetilde{\gamma}} \left(\theta - \int \vartheta_{\bar{h}} e^{\lambda}\, d\tau \right) \tag{7.23}$$

where $\widetilde{\gamma}$ is the projection of γ to the reduced phase space. We conclude that the reduced Hamiltonian conjugated to the time variable τ must be given by

$$\widetilde{H} = \int \vartheta_{\bar{h}} e^{\lambda} \tag{7.24}$$

Although it should follow just from consistency that this $\widetilde{H}$ is a well-defined function on $\widetilde{\mathcal{P}}$, we will show it explicitly. Before presenting a technical proof, let us give a physical interpretation for this Hamiltonian, which should by itself clarify why $\widetilde{H}$ is well-defined on the reduced phase space. From (6.42) and (6.16) we have

$$\widetilde{H} = \int \vartheta_{e^{\lambda} \bar{h}} = \int \vartheta_{\Psi_*^{-1} h} \tag{7.25}$$

since ϑ is the covariant volume element, it transforms nicely under a diffeomorphism, i.e., $\vartheta_{\Psi_*^{-1}h} = \Psi_*^{-1}\vartheta_h$. Using the invariance of the integral under automorphisms, we get

$$\widetilde{H} = \int \vartheta_h \tag{7.26}$$

That is, $\widetilde{H}$ is equal to the *area of the CMC slice* with mean curvature $K = -\tau$. Since this is a gauge-invariant property of the spacetime, it should be a well-defined function on the reduced phase space. This Hamiltonian associated with evolution along CMC slices is known as the *York Hamiltonian* [40].

7.3 The Hamiltonian is well-defined

Now, just for completeness, we will prove explicitly that $\widetilde{H}$ in (7.24) is a well-defined function on $\widetilde{\mathcal{P}}$. First we note that $\widetilde{H}$ is a well-defined function on $\hat{\mathcal{S}}$, $\widetilde{H}(\psi, \mathring{\sigma})$, since λ is the unique solution of (6.19) where σ^{ab} is determined by $\mathring{\sigma}$ via the mode expansions in (6.53) and (6.56) and the boundary value of λ is determined by $\psi = \Psi|_{\partial D}$ via (6.18). We therefore just need to show that $\widetilde{H}$ projects nicely through J to $\widetilde{\mathcal{P}}$, i.e., that $\widetilde{H}(\psi, \mathring{\sigma}) = \widetilde{H}(\psi', \mathring{\sigma}')$ whenever $J(\psi, \mathring{\sigma}) = J(\psi', \mathring{\sigma}')$.

Let us go back, temporarily, to the notation used in the theorem proven in point (i) after (6.72). We wish to have an explicit characterization of the pre-image under J of points in $T^*(G/H)$; in particular, we want to know how the gauge group H acts on $G \times \mathring{\mathfrak{g}}^*$, for the orbits of H are precisely the pre-images under J of points in

$T^*(G/H)$. Consider a transformation on $G \times \mathring{\mathfrak{g}}^*$, $(g, \sigma) \mapsto (g', \sigma')$, such that

$$J(g, \sigma) = J(g', \sigma') \tag{7.27}$$

This requires that $[g] = [g']$, and so there must exist $h \in H$ such that $g' = gh$. Also $J(g, \sigma)(V) = J(gh, \sigma')(V)$ for all vectors V at $[g] \in G/H$. This implies

$$\sigma(\Xi(X)) = \sigma'(\Xi(X')) \tag{7.28}$$

where X and X' are vectors at $g \in G$ and $gh \in G$, respectively, such that $q_* X = q_* X' = V$. Note that, given X at g, we can always choose $X' = r_{h*} X$ as the vector at gh. This follows from the fact that $q(g) = q(gh) = q \circ r_h(g)$, so $q_* X' = q_*(r_{h*} X) = q_* X = V$. Thus we have

$$\sigma(l_{g^{-1}*} X) = \sigma'(l_{(gh)^{-1}*} r_{h*} X) = \sigma'(l_{h^{-1}*} l_{g^{-1}*} r_{h*} X) = \sigma'(\mathrm{Ad}_{h^{-1}*} l_{g^{-1}*} X) \tag{7.29}$$

As we can choose any V at $[g]$, in particular we can choose $V = q_* l_{g*} \xi$, for any $\xi \in \mathfrak{g}$, so we conclude that $\sigma' = \mathrm{Ad}_h^* \sigma$ or, using the definition of the coadjoint action introduced in (6.87),

$$\sigma' = \mathrm{coad}_{h^{-1}} \sigma \tag{7.30}$$

Thus we can define the (right) H-action Γ_h on $G \times \mathring{\mathfrak{g}}^*$ by

$$\Gamma_h(g, \sigma) := (gh, \mathrm{coad}_{h^{-1}}\sigma) \tag{7.31}$$

and this is precisely the gauge group of $G \times \mathring{\mathfrak{g}}^*$, that is, $J(g, \sigma) = J(g', \sigma')$ if and only if $(g', \sigma') = \Gamma_h(g, \sigma)$ for some $h \in H$. Note that this result can be straightforwardly used to derive the kernel of J_*, as characterized in (6.93).

Particularizing to $\mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$, we consider the transformation

$$(\psi, \mathring{\sigma}) \mapsto (\psi', \mathring{\sigma}') = (\psi\varphi, \mathrm{coad}_{\varphi^{-1}}\mathring{\sigma}) \tag{7.32}$$

where $\varphi \in PSL(2, \mathbb{R})$. We are interested in computing how λ changes under this transformation. Recall that λ is the solution of

$$\bar{\nabla}^2\lambda + e^{-\lambda}\bar{\sigma}^{ab}\bar{\sigma}_{ab} - e^{\lambda}\chi = 0 \tag{7.33}$$

with boundary conditions

$$e^{\lambda}\bar{h}\Big|_{\partial D} = \psi_*^{-1}\gamma \tag{7.34}$$

The transformed λ, λ', will be the solution of

$$\bar{\nabla}^2\lambda' + e^{-\lambda'}\bar{\sigma}'^{ab}\bar{\sigma}'_{ab} - e^{\lambda'}\chi = 0 \tag{7.35}$$

with boundary conditions

$$e^{\lambda'}\bar{h}\Big|_{\partial D} = \psi'^{-1}_{*}\gamma \tag{7.36}$$

Let us first study how the boundary conditions for λ and λ' are related. Recall that $PSL(2,\mathbb{R})$ action on S^1 is simply the boundary action of the diffeomorphism part of the conformal isometries of D, that is, given $\varphi \in PSL(2,\mathbb{R}) \subset \mathit{Diff}^{+}(S^1)$, there exists a unique $\Phi \in \mathrm{Diff}^{+}(D)$ satisfying $\Phi|_{\partial D} = \varphi$ and a unique $\Omega \in C^{\infty}(D, \mathbb{R}^{+})$ such that

$$\bar{h} = \Phi_{*}\Omega\bar{h} \tag{7.37}$$

We have

$$e^{\lambda'}\bar{h}\Big|_{\partial D} = \varphi_{*}^{-1}\psi_{*}^{-1}\gamma = \varphi_{*}^{-1}\, e^{\lambda}\bar{h}\Big|_{\partial D} \tag{7.38}$$

so

$$\bar{h}\big|_{\partial D} = e^{-\lambda}\varphi_{*}\, e^{\lambda'}\bar{h}\Big|_{\partial D} = \varphi_{*}\, e^{-\lambda\circ\varphi}e^{\lambda'}\bar{h}\Big|_{\partial D} \tag{7.39}$$

so we can identify

$$e^{-\lambda\circ\varphi+\lambda'}\Big|_{\partial D} = \Omega\Big|_{\partial D} \tag{7.40}$$

which gives

$$\lambda'\Big|_{\partial D} = (\lambda\circ\varphi + \log\Omega)\,\Big|_{\partial D} \tag{7.41}$$

This suggests a convenient change of variables from λ' to $\widetilde{\lambda}$ defined as

$$\widetilde{\lambda} := \Phi_{*}(\lambda' - \log\Omega) \tag{7.42}$$

since this implies that $\widetilde{\lambda}$ is subjected to the same boundary conditions as λ, that is,

$$\widetilde{\lambda}\Big|_{\partial D} = \lambda\Big|_{\partial D} \tag{7.43}$$

It is natural to wonder if $\widetilde{\lambda}$ also satisfies the same equation as λ, which would imply that $\widetilde{\lambda} = \lambda$. To investigate this possibility, consider the action of the Lichnerowicz operator (for $\mathring{\sigma}$) on $\widetilde{\lambda}$,

$$L\widetilde{\lambda} := \bar{\nabla}^2\widetilde{\lambda} + e^{-\widetilde{\lambda}}\bar{\sigma}^{ab}\bar{\sigma}_{ab} - e^{\widetilde{\lambda}}\chi \tag{7.44}$$

The first term can be manipulated as

$$\nabla^2_{(\bar{h})}\widetilde{\lambda} = \nabla^2_{(\bar{h})}\Phi_*(\lambda' - \log\Omega) = \Phi_*\left[\nabla^2_{(\Phi_*^{-1}\bar{h})}(\lambda' - \log\Omega)\right] = \Phi_*\left[\nabla^2_{(\Omega\bar{h})}(\lambda' - \log\Omega)\right] \tag{7.45}$$

where $\nabla^2_{(h)}$ denotes Laplacian covariantly associated with h, and it transforms in a simple manner under the Weyl scaling

$$\nabla^2_{(\bar{h})}\widetilde{\lambda} = \Phi_*\left[\Omega^{-1}\nabla^2_{(\bar{h})}(\lambda' - \log\Omega)\right] = \Phi_*\left[\Omega^{-1}\nabla^2_{(\bar{h})}\lambda'\right] \tag{7.46}$$

in which we used that $\nabla^2_{(\bar{h})}\log\Omega = 0$.[33] The second term gives,

$$e^{-\widetilde{\lambda}}\bar{\sigma}^{ab}\bar{\sigma}_{ab} = e^{-\Phi_*(\lambda' - \log\Omega)}\bar{\sigma}^{ab}\bar{\sigma}_{ab} = \Phi_*\left[e^{-\lambda'}\Omega\,\Phi_*^{-1}\Big(\bar{\sigma}^{ab}\bar{\sigma}_{ab}\Big)\right] \tag{7.47}$$

[33] As explained before, this follows from the fact that $\Phi_*^{-1}\bar{h}$ and $\Omega\bar{h}$ are both flat and the Ricci curvature transforms like $R_{(\Omega\bar{h})} = \Omega^{-1}\Big(R_{(\bar{h})} - \nabla^2_{(\bar{h})}\log\Omega\Big)$.

and the $\bar{\sigma}^{ab}\bar{\sigma}_{ab}$ term can be written as

$$\Phi_*^{-1}\Big(\bar{\sigma}^{ab}\bar{\sigma}_{ab}\Big) = \Phi_*^{-1}\Big(\bar{h}_{ac}\bar{h}_{bd}\bar{\sigma}^{ab}\bar{\sigma}^{cd}\Big) = \Omega\bar{h}_{ac}\Omega\bar{h}_{bd}(\Phi_*^{-1}\bar{\sigma}^{ab})(\Phi_*^{-1}\bar{\sigma}^{cd}) \tag{7.48}$$

and so

$$e^{-\widetilde{\lambda}}\bar{\sigma}^{ab}\bar{\sigma}_{ab} = \Phi_*\left[\Omega^3 e^{-\lambda'}\bar{h}_{ac}\bar{h}_{bd}(\Phi_*^{-1}\bar{\sigma}^{ab})(\Phi_*^{-1}\bar{\sigma}^{cd})\right] \tag{7.49}$$

The third term is easy to compute,

$$e^{\widetilde{\lambda}}\chi = e^{\Phi_*(\lambda' - \log\Omega)}\chi = \Phi_*\left[\Omega^{-1}e^{\lambda'}\chi\right] \tag{7.50}$$

Now let us see how the transformation $\mathring{\sigma} \mapsto \mathring{\sigma}'$ is expressed as $\bar{\sigma}^{ab} \mapsto \bar{\sigma}'^{ab}$. Given a vector field $\hat{\xi}$ on S^1, representing an element of the algebra of *Diff*$^+(S^1)$, we have

$$\mathring{\sigma}'(\hat{\xi}) = \mathrm{Ad}^*_\varphi\mathring{\sigma}(\hat{\xi}) = \mathring{\sigma}(\mathrm{Ad}_{\varphi*}\hat{\xi}) = \mathring{\sigma}(\varphi_*\hat{\xi}) \tag{7.51}$$

where we used that $\mathrm{ad}_\varphi\hat{\xi} = \varphi_*\hat{\xi}$. This formula for the adjoint action on a group of diffeomorphisms is intuitive because ad_φ is a linear map on the space of vectors fields $\hat{\xi}$ and it carries a representation of the group, so it is natural to guess that it should act as the push-forward φ_*.[34] From the identification (6.60) we obtain

$$\int \vartheta_{\bar{h}}\, \bar{\sigma}'^{ab} \pounds_\xi \bar{h}_{ab} = \int \vartheta_{\bar{h}}\, \bar{\sigma}^{ab} \pounds_{\Phi_*\xi} \bar{h}_{ab} \tag{7.52}$$

where ξ is any extension of $\hat{\xi}$ to D. In principle this equality is also valid for an arbitrary extension of φ to D, but we chose to extend it as Φ for convenience. Pulling out the diffeomorphism on the right side of the equality gives

$$\int \vartheta_{\bar{h}}\, \bar{\sigma}^{ab} \pounds_{\Phi_*\xi} \bar{h}_{ab} = \int \Phi_* \left[\vartheta_{\Phi_*^{-1}\bar{h}}\, (\Phi_*^{-1}\bar{\sigma}^{ab}) \pounds_\xi (\Phi_*^{-1}\bar{h}_{ab})\right] = \int \vartheta_{\Omega\bar{h}}\, (\Phi_*^{-1}\bar{\sigma}^{ab}) \pounds_\xi (\Omega \bar{h}_{ab}) \tag{7.53}$$

where we used the invariance of the integral under a diffeomorphism. Note that

$$(\Phi_*^{-1}\bar{\sigma}^{ab}) \pounds_\xi (\Omega \bar{h}_{ab}) = (\Phi_*^{-1}\bar{\sigma}^{ab}) \left[\pounds_\xi \Omega \bar{h}_{ab} + \Omega \pounds_\xi \bar{h}_{ab}\right] = \Omega (\Phi_*^{-1}\bar{\sigma}^{ab}) \pounds_\xi \bar{h}_{ab} \tag{7.54}$$

where the first term inside the brackets vanishes because of the transverseness of $\bar{\sigma}^{ab}$, i.e., $(\Phi_*^{-1}\bar{\sigma}^{ab})\bar{h}_{ab} = \Omega^{-1}\Phi_*^{-1}[\bar{\sigma}^{ab}\bar{h}_{ab}] = 0$. Thus,

$$\int \vartheta_{\bar{h}}\, \bar{\sigma}'^{ab} \pounds_\xi \bar{h}_{ab} = \int \vartheta_{\bar{h}}\, (\Omega^2 \Phi_*^{-1}\bar{\sigma}^{ab}) \pounds_\xi \bar{h}_{ab} \tag{7.55}$$

[34]Here is the proof. Let $\mathcal{M}$ be a manifold and Diff($\mathcal{M}$) be its group of diffeomorphisms. Let the action of Diff($\mathcal{M}$) on $\mathcal{M}$ be denoted by $\Gamma_p(\psi) := \psi(p)$, where $p \in \mathcal{M}$, so the vector field induced by $\xi \in \mathfrak{diff}(\mathcal{M})$ is given by $X_p^\xi := \Gamma_{p*}\xi$. Given two diffeomorphisms ϕ and ψ, note that $\psi \circ \phi = \psi\phi\psi^{-1}\psi = \mathrm{Ad}_\psi \circ \psi$, which can also be expressed as $\Gamma_{\psi(p)} \circ \mathrm{Ad}_\psi = \psi \circ \Gamma_p$. Taking the derivative of this expression and acting on ξ gives $X_{\psi(p)}^{\mathrm{Ad}_\psi \xi} = \psi_*(X_p^\xi)$, where $\mathrm{ad}_\psi := \mathrm{Ad}_{\psi*}$. Under the identification $\xi \sim X^\xi$ this reads $\mathrm{ad}_\psi \xi = \psi_* \xi$.

so we can identify

$$\bar{\sigma}'^{ab} = \Omega^2 \Phi_*^{-1} \bar{\sigma}^{ab} \tag{7.56}$$

This identification is legitimate because $\Omega^2 \Phi_*^{-1} \bar{\sigma}^{ab}$ is traceless and transverse with respect to $\Phi_* \Omega \bar{h} = \bar{h}$ and it is associated with $\mathring{\sigma}'$ via (6.60). With this expression, we can rewrite the second term in (7.44) as

$$e^{-\widetilde{\lambda}} \bar{\sigma}^{ab} \bar{\sigma}_{ab} = e^{-\Phi_*(\lambda' - \log \Omega)} \bar{\sigma}^{ab} \bar{\sigma}_{ab} = \Phi_* \left[\Omega^{-1} e^{-\lambda'} \bar{\sigma}'^{ab} \bar{\sigma}'_{ab} \right] \tag{7.57}$$

Putting the three terms together we obtain

$$L\widetilde{\lambda} = \Phi_* \Omega^{-1} \left[\bar{\nabla}^2 \lambda' + e^{-\lambda'} \bar{\sigma}'^{ab} \bar{\sigma}'_{ab} - e^{\lambda'} \chi \right] = 0 \tag{7.58}$$

so we conclude that $\widetilde{\lambda}$ is in fact equal to λ, which implies that

$$\lambda' = \lambda \circ \Phi + \log \Omega \tag{7.59}$$

The transformed Hamiltonian is then

$$\widetilde{H}(\psi', \mathring{\sigma}') = \int \vartheta_{\bar{h}} e^{\lambda'} = \int \vartheta_{\bar{h}} \Omega e^{\Phi_*^{-1} \lambda} = \int \vartheta_{\Phi_*^{-1} \bar{h}} \Phi_*^{-1} e^{\lambda} = \int \Phi_*^{-1} \left[\vartheta_{\bar{h}} e^{\lambda} \right] = \widetilde{H}(\psi, \mathring{\sigma}) \tag{7.60}$$

finishing the proof that $\widetilde{H}$ is indeed a well-defined function on $\widetilde{\mathcal{P}}$.

8 Approximations for the Hamiltonian

The Hamiltonian generating evolution along CMC slices is, as we have just seen, a τ-dependent function on the reduced phase space given by the area of the CMC slice with mean curvature $K = -\tau$. Despite its simple geometric interpretation, this function is extremely complicated: in order to evaluate $\widetilde{H}$ at a point $p \in \widetilde{\mathcal{P}}$, we need to take any $(\psi, \mathring{\sigma}) \in J^{-1}(p)$ and solve the differential equation (6.19) for λ, in which $\mathring{\sigma}$ enters through the term $\bar{\sigma}^{ab}\bar{\sigma}_{ab}$ and ψ enters as a boundary condition, and then integrate e^{λ} over the disc D. It is thus worthwhile to investigate whether there are certain regimes in which the Hamiltonian can be approximated by something simpler. There are three independent length scales in our problem,

- the boundary length, ℓ;
- the AdS length, $\ell_{\text{AdS}} := \frac{1}{\sqrt{-\Lambda}}$;
- the Planck length, $\ell_P := \hbar G$;

and we recall that we are considering $c = 1$. As usual in quantum field theory, we shall also consider $\hbar = 1$, so $\ell_P = G$. We wish to explore the different limits of these scales.

8.1 Reintroducing the physical scales

Let us explicitly reintroduce these scales in our formulas. As we wish to take $\bar{h}_{ab}$ as corresponding to the unit-radius (dimensionless) disc, $\bar{h} = dr^2 + r^2 d\theta^2$, let us redefine the conformal map in (6.16) as

$$h_{ab} = \left(\frac{\ell}{2\pi}\right)^2 \Psi_* e^{\lambda} \bar{h}_{ab} \tag{8.1}$$

and the boundary condition (6.18) reads

$$e^{\lambda}\bar{h}\Big|_{\partial D} = \left(\frac{\ell}{2\pi}\right)^{-2} \psi_*^{-1}\gamma \tag{8.2}$$

In this way, the length scale is fully encoded in h_{ab}, while ψ and λ are dimensionless. In particular, if $\gamma = (\ell/2\pi)^2 d\theta^2$, then

$$e^{\lambda} d\theta^2\Big|_{\partial D} = \psi_*^{-1} d\theta^2 \tag{8.3}$$

There is no reason not to make that choice since γ can always be parametrized by proper length, and we can define the unit disc by a uniform rescaling of the boundary. Similarly, a reasonable redefinition of the map in (6.17) would be

$$\sigma^{ab} = \left(\frac{\ell}{2\pi}\right)^{-3} \Psi_* e^{-2\lambda} \bar{\sigma}^{ab} \tag{8.4}$$

so that $\bar{\sigma}^{ab}$ is dimensionless, since σ^{ab} has dimensions of *length*$^{-3}$.[35] However, a more natural redefinition turns out to be

$$\sigma^{ab} = 16\pi\ell_P \left(\frac{\ell}{2\pi}\right)^{-4} \Psi_* e^{-2\lambda}\bar{\sigma}^{ab} \tag{8.5}$$

which also leads to a dimensionless $\bar{\sigma}^{ab}$. The reason for introducing this $\sim G$ factor is because the correct expression for the conjugate momentum, when the overall $(16\pi G)^{-1}$ factor in the action is not omitted, is

$$\pi^{ab} = \frac{1}{16\pi G}\sqrt{\det(h)}\left(\sigma^{ab} + \frac{1}{2}\tau h^{ab}\right) \tag{8.6}$$

so this $(16\pi G)^{-1}$ factor appears in the rightmost side of the expression (6.30) for the pre-symplectic form Ω. In this way, the inclusion of the $16\pi\ell_P$ in (6.17) exactly cancels this factor in the corresponding (pre-)symplectic form on $\bar{\mathcal{S}}$, $\bar{\omega}$, given in (6.44). In fact, (6.44) is correct as it stands, as ℓ will not appear either. This is because ℓ,

[35] A brief comment on how we are assigning dimensions to tensors. The length of a curve γ is given by $\int dt\sqrt{h_{ab}\dot{\gamma}^a\dot{\gamma}^b}$, and this must have dimension L. The tangent vector $\dot{\gamma}$ can be defined for its action of functions as $\dot{\gamma}(f) = \frac{d}{dt}f(\gamma(t))$, so we see that it is natural to associated dimension T^{-1} to $\dot{\gamma}$, where T is the dimension of the parameter t. Therefore, from the formula for the length of γ, we have $L \sim T\sqrt{h_{ab}T^{-1}T^{-1}}$, so we conclude that the metric h_{ab} must have dimension L^2. This is consistent with a typical expression like $h = dx^2 + dy^2$, where the coordinates x and y have dimension L. The extrinsic curvature is defined by a formula like $K_{ab} = \frac{1}{2}\pounds_n h_{ab}$ where n is a unit vector (and so n must have dimension L^{-1}). Therefore, $K_{ab} \sim L^{-1}L^2 \sim L$. Since $h^{ab}h_{bc} = \delta^a_c$, and δ^a_c is naturally dimensionless, we must have $h^{ab} \sim L^{-2}$. Note that this gives $K = K_{ab}h^{ab} \sim LL^{-2} \sim L^{-1}$, which is consistent with the interpretation that K gives the rate (per unit of normal length) that the volume $\delta\vartheta$ of a piece of the surface changes, $K = \frac{1}{\delta\vartheta}\frac{d\delta\vartheta}{dl}$. Lastly, $K^{ab} = h^{ac}h^{bd}K_{cd} \sim L^{-2}L^{-2}L \sim L^{-3}$, so σ^{ab} have dimension L^{-3}.

as introduced above, can be completely absorbed in an effective redefinition of λ,

$$h_{ab} = \Psi_* e^{\lambda+2\log(\ell/2\pi)} \bar{h}_{ab}$$

$$\sigma^{ab} = 16\pi\ell_P \Psi_* e^{-2(\lambda+2\log(\ell/2\pi))} \bar{\sigma}^{ab} \tag{8.7}$$

and λ does not appear explicitly in (6.44). This means that all formulas in the remaining of that section are correct as they stand. In particular, the association $\bar{\sigma}^{ab} \leftrightarrow \mathring{\sigma}$, the projection map J, and the fact that the symplectic form on $\widetilde{\mathcal{P}}$ is equal to the canonical 2-form associated with its cotangent bundle structure all remain unchanged. There are only two formulas that need to be updated in view of (8.7), the Lichnerowicz equation (6.19) and the reduced Hamiltonian (7.24). Since the Lichnerowicz equation is a direct consequence of the Hamiltonian constraint, (6.11), which is independent of the overall factor $(16\pi G)^{-1}$ in the action, we have

$$\bar{\nabla}^2 \left(\lambda + 2\log(\ell/2\pi)\right) + e^{-(\lambda+2\log(\ell/2\pi))} \left(16\pi\ell_P \bar{\sigma}^{ab}\right)\left(16\pi\ell_P \bar{\sigma}_{ab}\right) - e^{(\lambda+2\log(\ell/2\pi))}\chi = 0 \tag{8.8}$$

which gives

$$\bar{\nabla}^2\lambda + \left(\frac{32\pi^2\ell_P}{\ell}\right)^2 \bar{\sigma}^{ab}\bar{\sigma}_{ab}\, e^{-\lambda} - \left(\frac{\ell}{2\pi}\right)^2 \left(\frac{2}{\ell_{AdS}^2} + \frac{\tau^2}{2}\right) e^{\lambda} = 0 \tag{8.9}$$

where we wrote χ explicitly to display the AdS length. The reduced Hamiltonian gets a factor $(16\pi G)^{-1}$ from the action, and another from the effective redefinition of λ, so it becomes

$$\widetilde{H} = (16\pi\ell_P)^{-1} \int \vartheta_{\bar{h}} e^{(\lambda + 2\log(\ell/2\pi))} = \frac{\ell^2}{64\pi^3 \ell_P} \int \vartheta_{\bar{h}} e^{\lambda} \tag{8.10}$$

Note that it has dimension of *length* because it is conjugated to a time, τ, which has dimension of *length*$^{-1}$. While the overall factor in the Hamiltonian is not particularly interesting, the different length scales still enter implicitly through λ, which is a solution of (8.9).

Let us now study more carefully the behavior of the Lichnerowicz equation (8.9) in different regimes. For reference, let us cast it in the more compact form

$$\bar{\nabla}^2 \lambda + \kappa e^{-\lambda} - \bar{\chi} e^{\lambda} = 0 \tag{8.11}$$

where

$$\kappa := \left(\frac{32\pi^2 \ell_P}{\ell} \right)^2 \bar{\sigma}^{ab} \bar{\sigma}_{ab} \tag{8.12}$$

is some sort of kinetic energy term, since it is quadratic in $\bar{\sigma}^{ab}$ which is a "momentum" variable, and

$$\bar{\chi} := \left(\frac{\ell}{2\pi} \right)^2 \left(\frac{2}{\ell_{AdS}^2} + \frac{\tau^2}{2} \right) \tag{8.13}$$

is a dimensionless version of χ, just a (time-dependent) constant. It is worth noting that this equation can be derived from a variational principle. The underlying functional is given by

$$I[\lambda] := \int_D \vartheta_{\bar{h}} \left(\frac{1}{2} |\bar{\nabla}\lambda|^2 + \kappa e^{-\lambda} + \bar{\chi} e^{\lambda} \right) \tag{8.14}$$

where $|\bar{\nabla}\lambda|^2 = \bar{h}^{ab}\bar{\nabla}_a\lambda\bar{\nabla}_b\lambda$. In the extremization, the variable $\lambda : D \to \mathbb{R}$ is varied with its boundary value fixed by (8.3). Computing the variation of I up to second order in $\delta\lambda$ gives

$$\delta I = -\int \vartheta_{\bar{h}} \left(\bar{\nabla}^2\lambda + \kappa e^{-\lambda} - \bar{\chi} e^{\lambda} \right) \delta\lambda + \frac{1}{2} \int \vartheta_{\bar{h}} \left(|\bar{\nabla}\delta\lambda|^2 + \kappa e^{-\lambda}\delta\lambda^2 + \bar{\chi} e^{\lambda}\delta\lambda^2 \right) \tag{8.15}$$

so from the first-order term we recover (8.11) and from the second-order term we conclude that the extremum is a (global) minimum of I. There is a simple pictorial interpretation for this variational principle: we imagine λ as specifying the configuration of a membrane in $D \times \mathbb{R}$. This membrane is a cross-section of this (solid) cylindrical space, fixedly attached to the boundary (as determined by ψ), which can be deformed up and down in the bulk. The first term in the functional is an *elastic potential energy* of a "Hookean" membrane[36] and, in the minimization process,

[36] The term "Hookean" is used here to describe an elastic membrane that sustains a tension proportional to the local stretching factor, i.e., $\tau = k(\delta a'/\delta a)$ where $\delta a'$ is the area of a piece of membrane with relaxed area δa. If $\delta a'$ is further stretched by an infinitesimal normal displacement δs, its area will change as $\delta a' \to \delta a' + \int dl \delta s$, where dl is the length element along the boundary of $\delta a'$. The work done to stretch this piece is $\delta U = \int dl\, \tau \delta s = \int dl\, k(\delta a'/\delta a)\delta s = (k/\delta a) \int d(\delta a')(\delta a') = \frac{k}{2}\delta a'^2/\delta a$. The total potential energy of the membrane is thus $U = \int \delta a \frac{k}{2} \left(\frac{\delta a'}{\delta a}\right)^2$. Now consider the membrane in the Euclidean cylinder $D \times \mathbb{R}$, and say that $z = \lambda = 0$ is the relaxed configuration. Then for a generic configuration $z = \lambda(x)$, obtained by deforming the membrane vertically, $(x, 0) \to (x, \lambda(x))$, a relaxed piece of area $\delta a = d^2x$ will be stretched to $\delta a' = \sqrt{1 + |\nabla\lambda|^2} d^2x$, so the potential energy (for $k = 1$) will be given by $U = \int_D d^2x \frac{1}{2}\left(1 + |\nabla\lambda|^2\right) = \frac{\pi}{2} + \int_D d^2x \frac{1}{2}|\nabla\lambda|^2$.

it offers a resistance against stretching the membrane too much; the second term, proportional to the *kinetic energy* κ, tries to pull the membrane upwards; and the third term, acting as a time-dependent *external force*, tries to pull it downwards. This picture provides a clear way to see that the Hamiltonian tends to increase with the kinetic energy κ, just as in typical systems.

8.2 Liouville equation

The Lichnerowicz equation, (8.11), is a quite complicated non-linear partial differential equation. It is remarkable that it can be solved exactly in the case of zero momentum. In 1853, Liouville decided to study the equation

$$\frac{\partial^2 \log \Omega}{\partial u \partial v} - \frac{\bar{\chi}}{4}\Omega = 0 \tag{8.16}$$

managing to find a general analytic solution for it [72]. This is equation, which we shall call the *Liouville equation*, is quite reminiscent of (8.11) when the kinetic term is absent,

$$\bar{\nabla}^2 \lambda - \bar{\chi} e^{\lambda} = 0 \tag{8.17}$$

In fact, they are related by a change of variables $\Omega = e^{\lambda}$ and

$$\begin{aligned} u &= x + iy \\ v &= x - iy \end{aligned} \tag{8.18}$$

where x and y are Cartesian coordinates on $(D, \bar{h})$. Since u and v are just the complex coordinates associated with (x, y), we shall use the more standard notation of $z = x + iy$ and $\bar{z} = x - iy$. The general solution for this equation is given by

$$\lambda = \log\left[\frac{8}{\bar{\chi}}\frac{f'(z)g'(\bar{z})}{(1 - f(z)g(\bar{z}))^2}\right] \tag{8.19}$$

where f and g are arbitrary meromorphic and anti-meromorphic functions on $\mathbb{D} \sim D$, and f' and g' are their respective first derivatives. Since λ must be smooth and real on D, we must impose some additional conditions for f and g. To ensure reality, the simplest condition is to take $g(\bar{z}) = \overline{f(z)}$, so we would get

$$\lambda = \log\left[\frac{8}{\bar{\chi}}\frac{|f'(z)|^2}{(1 - |f(z)|^2)^2}\right] \tag{8.20}$$

To ensure smoothness, we may require that f' is nowhere vanishing on $\mathbb{D}$ and that the image of f is either entirely in $\mathbb{D}$ or in $\mathbb{C} - \mathbb{D}$, so the denominator never vanishes.

An interesting solution is the one corresponding to the *symmetric diamond*, i.e., the state $(\psi, \mathring{\sigma}) = (\mathrm{I}, 0)$. This is the diamond whose maximal surface is an Euclidean disc. This case is particularly simple so that we do not even need to use the general solution described above, as the corresponding Lichnerowicz reduces to an ordinary differential equation for $\lambda_0(r, \theta) = \lambda_0(r)$, in polar coordinates,

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{\lambda_0}{dr}\right) - \bar{\chi}e^{\lambda_0} = 0 \tag{8.21}$$

satisfying boundary conditions $\lambda(1) = 0$ and $\lambda_0'(0) = 0$ (where the derivative condition is so that $\lambda_0(r, \theta)$ is regular at the origin). The solution is given by

$$\lambda_0(r, \theta) = \log\left(\frac{8C}{(\bar{\chi}C - r^2)^2}\right) , \quad C = \frac{\bar{\chi} + 4 + \sqrt{8\bar{\chi} + 16}}{\bar{\chi}^2} \tag{8.22}$$

Notice that this result can also be obtained from the general solution (8.20) for the function

$$f(z) = \frac{z}{\sqrt{\bar{\chi}C}} \tag{8.23}$$

In fact, the choice $f(z) = \sqrt{\bar{\chi}C}/z$ also leads to the same result, revealing that different fs may correspond to the same solution.

We can compute explicitly the Hamiltonian for the symmetric diamond. Inserting (8.22) into (8.10) we get

$$\widetilde{H}_0 = \frac{\ell^2}{64\pi^3 \ell_P} 2\pi \int_0^1 r dr \, e^{\lambda_0(r)} = \frac{\ell^2}{32\pi^2 \ell_P \left(1 + \sqrt{\bar{\chi}/2 + 1}\right)} \tag{8.24}$$

Observe that, as expected, the Hamiltonian is maximal at $\tau = 0$, decreasing monotonically as τ^2 increases.

8.3 Large boundary length

An interesting regime is that of a large boundary length compared to the AdS_3 length, i.e., $\ell \gg \ell_{AdS}$. Basically we are interested in the limit where $\bar{\chi}$ is large, presumably when compared to the scales of the kinetic term κe^{λ} and the boundary values $\lambda|_\partial = 2\log\psi'$, but the precise conditions for the approximations to hold will be defined later. Note that this is also achieved for very late or earlier times, when $|\tau| \gg \ell^{-1}$. In this limit we expect that the contribution of the kinetic term should be subdominant when compared to the "external force" $\bar{\chi}e^{\lambda}$. Thus, it should in principle be possible to perturbatively expand the Hamiltonian around the zero-momentum solution, which we know exactly. Instead of tackling the full problem, let us consider a simpler version of it where we expand around the symmetric diamond state $(\psi, \mathring{\sigma}) = (\mathrm{I}, 0)$.

Let us assume that, in a certain neighborhood of the state $(\psi, \mathring{\sigma}) = (\mathrm{I}, 0)$, the solution to the Lichnerowicz equation (8.11) can be written as

$$\lambda = \lambda_0 + \delta\lambda \tag{8.25}$$

where λ_0 is the solution associated with the $(\mathrm{I}, 0)$, given in (8.22), and $\delta\lambda \ll 1$ is a small correction. The Lichnerowicz equation then reads

$$\bar{\nabla}^2(\lambda_0 + \delta\lambda) + \kappa e^{-\lambda_0} e^{-\delta\lambda} - \bar{\chi} e^{\lambda_0} e^{\delta\lambda} = 0 \tag{8.26}$$

which can be approximated, to first order in $\delta\lambda$, as

$$\bar{\nabla}^2 \delta\lambda - \bar{\chi} e^{\lambda_0} \delta\lambda = -\kappa e^{-\lambda_0} \tag{8.27}$$

where we expanded $e^{\delta\lambda} \approx 1 + \delta\lambda$ and neglected the (higher order) term $\kappa\delta\lambda$. Since $\lambda_0|_\partial = 0$, the boundary condition for λ must be entirely supplied by $\delta\lambda|_\partial = 2\log\psi'$. Since $\delta\lambda \ll 1$ we must require

$$2\log\psi' \ll 1 \tag{8.28}$$

that is, the diffeomorphism must be sufficiently close to the identity. Notice that if $\kappa = 0$ and $\psi' = 1$ the perturbation would vanish and the solution would be $\lambda = \lambda_0$, as expected. It is convenient to decompose the perturbation into two parts depending on the "source" causing $\delta\lambda$ not to vanish. Namely, write $\delta\lambda = \delta\lambda_q + \delta\lambda_p$, where $\delta\lambda_q$ and $\delta\lambda_p$ are defined by the equations

$$\bar{\nabla}^2\delta\lambda_q - \bar{\chi}e^{\lambda_0}\delta\lambda_q = 0\,, \qquad \delta\lambda_q|_\partial = 2\log\psi' \tag{8.29}$$

$$\bar{\nabla}^2\delta\lambda_p - \bar{\chi}e^{\lambda_0}\delta\lambda_p = -\kappa e^{-\lambda_0}\,, \quad \delta\lambda_p|_\partial = 0 \tag{8.30}$$

In this way, $\delta\lambda_q$ satisfies a homogeneous equation "sourced" by a nontrivial boundary condition (so it is associated with "position" variables $q = \psi$) and $\delta\lambda_p$ satisfies an inhomogeneous equation, "sourced" by the kinetic term, with a trivial boundary condition (so it is associated with "momentum" variables $p = \mathring{\sigma}$). Let us analyze the typical behaviors of these two equations below.

The position-sourced perturbation equation, (8.29), can be cast as a minimization problem associated with the functional

$$I_q[u] = \frac{1}{2}\int d^2x \left(|\bar{\nabla}u|^2 + \bar{\chi}e^{\lambda_0}u^2\right) \tag{8.31}$$

The minimizer $u : D \to \mathbb{R}$ of I_q, satisfying the boundary condition $u|_\partial = 2\log\psi'$, is the solution $\delta\lambda_q$ of (8.29). Using the "elastic membrane" interpretation discussed around (8.14), we see that the term $|\bar{\nabla}u|^2$ tries to keep the membrane as flat as possible and term u^2 tried to keep it as close to $u = 0$ as possible. The u^2 term comes accompanied by the factor $\bar{\chi}e^{\lambda_0}$. In the limit $\bar{\chi} \gg 1$ we can approximate C in (8.22) by $(\bar{\chi} + \sqrt{8\bar{\chi}})/\bar{\chi}^2$, so we have

$$\bar{\chi}e^{\lambda_0} \approx \frac{8}{\left(1 - r^2 + \sqrt{8/\bar{\chi}}\right)^2} \tag{8.32}$$

Note that this factor is ~ 10 in most of the interior of the disk, growing to a huge value $\sim \chi$ near the boundary. This means that, if $\bar{\chi}$ is large enough, it will bring the membrane to zero as soon as it leaves the boundary, which would then just remain close to zero in the interior. See an example in Fig. 6.

The momentum-sourced perturbation equation, (8.30), can also be cast as a minimization problem, associated with the functional

$$I_p[u] = \frac{1}{2}\int d^2x \left(|\bar{\nabla}u|^2 + \bar{\chi}e^{\lambda_0}u^2 - 2\kappa e^{-\lambda_0}u\right) \tag{8.33}$$

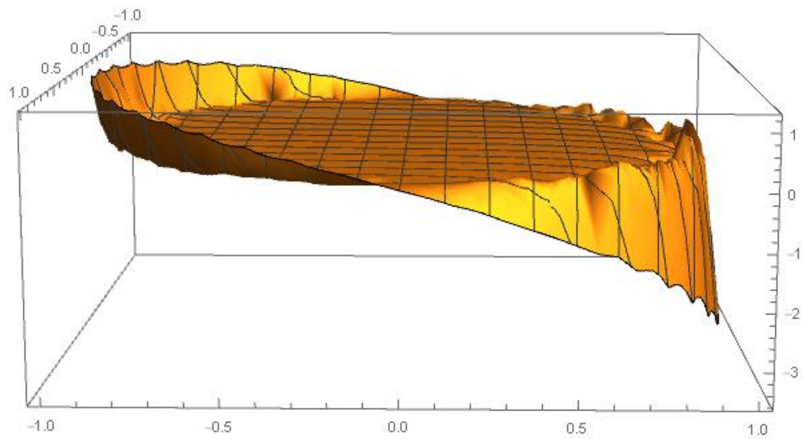


Figure 6: Typical solution for $\delta\lambda_q$ in the large $\bar{\chi}$ limit.

The minimizer, satisfying the boundary condition $u|_\partial = 0$, is the solution $\delta\lambda_p$ of (8.30). In the membrane interpretation, the terms $|\bar{\nabla}u|^2$and u^2 act like before, and the term linear in u tries to pull the membrane up. The coefficient of the linear term is (-2 times)

$$\kappa e^{-\lambda_0} \approx \frac{\kappa\bar{\chi}}{8}\left(1 - r^2 + \sqrt{8/\bar{\chi}}\right)^2 \tag{8.34}$$

Note that this factor is of order $\sim \kappa\bar{\chi}/8$ in most of the interior of the disc, going to $\sim \kappa$ near the boundary. Therefore the kinetic term may actually have a significant effect near the center of the disc. To estimate its influence, we combine the u^2 and the u terms by "completing the square",

$$\bar{\chi}e^{\lambda_0}u^2 - 2\kappa e^{-\lambda_0}u = \bar{\chi}e^{\lambda_0}\left(u - \frac{\kappa e^{-\lambda_0}}{\bar{\chi}e^{\lambda_0}}\right)^2 - \frac{\kappa^2}{\bar{\chi}}e^{-3\lambda_0} \tag{8.35}$$

Hence we see that the effect of these terms combined in the functional is to pull the membrane toward the “equilibrium” configuration $u_{\rm eq}$,

$$u_{\rm eq} := \frac{\kappa e^{-\lambda_0}}{\bar{\chi} e^{\lambda_0}} = 4\pi^2 \ell_P^2 \chi \bar{\sigma}^2 \left(1 - r^2 + \sqrt{8/\bar{\chi}}\right)^4 \tag{8.36}$$

where we used the definition (8.12) of κ in terms of $\bar{\sigma}^2 := \bar{\sigma}^{ab}\bar{\sigma}_{ab}$ and the definition (8.13) of $\bar{\chi}$ in terms of χ. The term $|\bar{\nabla} u|^2$ will offer some resistance to deforming membrane from $u = 0$ (the flattest configuration) to $u_{\rm eq}$, so we expect that u will be (roughly) bounded by $u_{\rm eq}$. Nonetheless, we still expect that the minimizer u will be of the same magnitude as $u_{\rm eq}$, so in order for our approximation to be legitimate (i.e., $u \ll 1$) we need to assume $u_{\rm eq} \ll 1$, which implies

$$\bar{\sigma}^2 \ll \frac{(\ell/\ell_P)^2}{16\pi^4 \bar{\chi}} \tag{8.37}$$

If, for example, τ^2 is not large so that $\chi \sim 2/\ell_{AdS}^2$, then the right-hand side becomes $\sim \ell_{AdS}^2/8\pi^2\ell_P^2$, which can be satisfied for a large range of $\mathring{\sigma}$ if $\ell_{AdS} \gg \ell_P$. Let us now consider the behavior of the solution near the boundary. There the coefficient $2\kappa e^{-\lambda_0}$ is suppressed by a factor of $\sim 8/\bar{\chi}$ while the coefficient $\bar{\chi} e^{\lambda_0}$ is enhanced and attains values of order $\sim \bar{\chi}$. Consequently, the equation tries to make the membrane very flat as it approaches the boundary. In particular, we expect the normal derivative of $\delta\lambda_p$ to be very small near the boundary.

Let us now proceed to find an approximated expression for the Hamiltonian $\widetilde{H}$ in this regime. It is interesting that we can obtain a general result for $\widetilde{H}$ without the need to fully solve (8.27). Expanding $e^{\delta\lambda} \approx 1 + \delta\lambda$ in (8.10) we get

$$\widetilde{H} \approx \widetilde{H}_0 + \frac{\ell^2}{64\pi^3 \ell_P} \int d^2x\, e^{\lambda_0} \delta\lambda \tag{8.38}$$

where $\widetilde{H}_0$ is given exactly in (8.24), but can also be approximated in the large $\bar{\chi}$ limit as

$$\widetilde{H}_0 \approx \frac{\ell^2}{16\pi^2 \ell_P \sqrt{2\bar{\chi}}} \tag{8.39}$$

Considering the decomposition of $\delta\lambda$ in $\delta\lambda_q$ and $\delta\lambda_p$, we can define two parts for the Hamiltonian perturbation,

$$\delta\widetilde{H}_q := \frac{\ell^2}{64\pi^3 \ell_P} \int d^2x\, e^{\lambda_0} \delta\lambda_q \tag{8.40}$$

$$\delta\widetilde{H}_p := \frac{\ell^2}{64\pi^3 \ell_P} \int d^2x\, e^{\lambda_0} \delta\lambda_p \tag{8.41}$$

so that $\widetilde{H} \approx \widetilde{H}_0 + \delta\widetilde{H}_q + \delta\widetilde{H}_p$. The slick observation here is that $\delta\lambda$ is integrated against a rotationally-symmetric function, e^{λ_0}, implying that only the zero Fourier mode of $\delta\lambda$ contributes to the Hamiltonian. Moreover, since $\delta\lambda$ satisfies a linear equation associated with the rotationally-symmetric operator $\bar{\nabla}^2 - \bar{\chi}e^{\lambda_0}$, the Fourier modes are all independent and we can therefore solve specifically for the zero mode.

Consider the zero modes of the "sources" given by

$$\Upsilon_0 := \frac{1}{2\pi}\int d\theta\, 2\log\psi'(\theta) \tag{8.42}$$

$$K_0(r) := \frac{1}{2\pi}\int d\theta\, \kappa(r,\theta)e^{-\lambda_0(r)} \tag{8.43}$$

and the zero modes of the perturbations

$$\delta\lambda_q^{(0)} := \frac{1}{2\pi}\int d\theta\, \delta\lambda_q \tag{8.44}$$

$$\delta\lambda_p^{(0)} := \frac{1}{2\pi}\int d\theta\, \delta\lambda_p \tag{8.45}$$

Integrating (8.29) and (8.30) in θ, using the polar expression for the Laplacian, we get

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\delta\lambda_q^{(0)}}{\partial r}\right) - \bar{\chi}e^{\lambda_0}\delta\lambda_q^{(0)} = 0\,, \qquad \delta\lambda_q^{(0)}|_\partial = \Upsilon_0 \tag{8.46}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\delta\lambda_p^{(0)}}{\partial r}\right) - \bar{\chi}e^{\lambda_0}\delta\lambda_p^{(0)} = -K_0\,, \quad \delta\lambda_p^{(0)}|_\partial = 0 \tag{8.47}$$

Note that the term proportional to $\bar{\chi}$ is precisely what appears in the integrands of (8.40) and (8.41). Therefore, if we integrate the equations above in rdr we should get

$$\delta\widetilde{H}_q = \frac{\ell^2}{64\pi^3\ell_P} 2\pi \int_0^1 rdr\, \frac{1}{\bar{\chi}} \left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\delta\lambda_q^{(0)}}{\partial r}\right)\right] \tag{8.48}$$

$$\delta\widetilde{H}_p = \frac{\ell^2}{64\pi^3\ell_P} 2\pi \int_0^1 rdr\, \frac{1}{\bar{\chi}} \left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\delta\lambda_p^{(0)}}{\partial r}\right) + K_0\right] \tag{8.49}$$

which gives

$$\delta\widetilde{H}_q = \frac{\ell^2}{32\pi^2\ell_P}\frac{1}{\bar{\chi}}\left.\frac{\partial\delta\lambda_q^{(0)}}{\partial r}\right|_{r=1} \tag{8.50}$$

$$\delta\widetilde{H}_p = \frac{\ell^2}{32\pi^2\ell_P}\left(\frac{1}{\bar{\chi}}\left.\frac{\partial\delta\lambda_p^{(0)}}{\partial r}\right|_{r=1} + \frac{1}{16\pi}\int_0^1 dr\, r(1-r^2)^2 \int d\theta\, \kappa\right) \tag{8.51}$$

where in the last integral, involving κ, we approximated $(1-r^2+\sqrt{8/\bar{\chi}})^2 \approx (1-r^2)^2$.

In our study of the typical behavior of $\delta\lambda_q$ and $\delta\lambda_p$, we noted that $\delta\lambda_q$ grows steeply near the boundary in the limit of large $\bar{\chi}$. In fact we can show that, in this limit,

$$\left.\frac{\partial\delta\lambda_q^{(0)}}{\partial r}\right|_{r=1} = \Upsilon_0\sqrt{\frac{\bar{\chi}}{2}} \tag{8.52}$$

The proof goes as follows. The function multiplying $\delta\lambda_q^{(0)}$ in (8.46) ca be approximated by

$$\bar{\chi}e^{\lambda_0} = \frac{8}{(1+\sqrt{8/\bar{\chi}}-r^2)^2} \approx \frac{8}{(1-\widetilde{r}^2)^2} \tag{8.53}$$

where

$$\widetilde{r} := \frac{r}{1+\sqrt{2/\bar{\chi}}} \tag{8.54}$$

In this new variable $\widetilde{r}$, equation (8.46) becomes approximately

$$\frac{1}{\widetilde{r}}\frac{d}{d\widetilde{r}}\left(\widetilde{r}\frac{d\delta\lambda_q^{(0)}}{d\widetilde{r}}\right) - \frac{8\delta\lambda_q^{(0)}}{(1-\widetilde{r}^2)^2} = 0 \tag{8.55}$$

whose general solution reads

$$\delta\lambda_q^{(0)} = \frac{c_1(1+\widetilde{r}^2) + c_2(2+\log\widetilde{r} - \widetilde{r}^2\log\widetilde{r})}{1-\widetilde{r}^2} \tag{8.56}$$

for real coefficients c_1 and c_2. In order for $\delta\lambda_q^{(0)}$ to be regular at $r = 0$ ($\widetilde{r} = 0$) it is necessary that $c_2 = 0$. The boundary condition is that $\delta\lambda_q^{(0)} = \Upsilon_0$ at $r = 1$, or $\widetilde{r} = 1/(1+\sqrt{2/\bar{\chi}})$, which implies

$$c_1 = \Upsilon_0\sqrt{\frac{2}{\bar{\chi}}} \tag{8.57}$$

The first derivative in r gives

$$\frac{d\delta\lambda_q^{(0)}}{dr} \approx \frac{d\delta\lambda_q^{(0)}}{d\widetilde{r}} = \Upsilon_0\sqrt{\frac{2}{\bar{\chi}}}\frac{4\widetilde{r}}{(1-\widetilde{r}^2)^2} \tag{8.58}$$

Note that at $r = 0$ we have $d\delta\lambda_q^{(0)}/dr = 0$, as expected (to avoid a kink at the origin). At $r = 1$, we obtain (8.52).

The above result yields,

$$\delta\widetilde{H}_q = \frac{\ell\ell_{AdS}}{16\pi\ell_P\sqrt{2\chi}}\Upsilon_0 \tag{8.59}$$

where we have written $\bar{\chi}$ in term of χ. We should note however that, to first order in $\delta\lambda$, Υ_0 actually vanishes. This follows from the fact that ψ is a diffeomorphism, so

$$2\pi = \int d\theta\, \psi'(\theta) = \int d\theta\, e^{\log\psi'} \approx \int d\theta\, (1 + \log\psi') = 2\pi + \pi\Upsilon_0 \tag{8.60}$$

Hence, in our approximations,

$$\delta\widetilde{H}_q = 0 \tag{8.61}$$

We have indicated that $\delta\lambda_p$ tends to flatness near the boundary, so the derivative term is subdominant in this case. Therefore we have

$$\delta\widetilde{H}_p = 2\pi\ell_P \int_0^1 dr\, r(1-r^2)^2 \int d\theta\, \bar{\sigma}^2 \tag{8.62}$$

Using the Fourier expansion of $\bar{\sigma}^{\mu\nu}$, given in (6.53), we obtain

$$\bar{\sigma}^2 := \bar{\sigma}^{\mu\nu}\bar{\sigma}_{\mu\nu} = \sum_{n,m\geq 0} 4r^{n+m}(a_n a_m + b_n b_m)\cos[(n-m)\theta] \tag{8.63}$$

whose zero mode is

$$\frac{1}{2\pi}\int d\theta\, \bar{\sigma}^2 = \sum_{n\geq 0} 4r^{2n}(a_n^2 + b_n^2) \tag{8.64}$$

Finally, using the result

$$\int_0^1 dr\, r^{2n+1}(1-r^2)^2 = \frac{1}{(n+1)(n+2)(n+3)} \tag{8.65}$$

we obtain

$$\delta\widetilde{H}_p = 16\pi^2\ell_P \sum_{n\geq 0} \frac{a_n^2 + b_n^2}{(n+1)(n+2)(n+3)} \tag{8.66}$$

Hence, the Hamiltonian in this regime becomes approximately

$$\widetilde{H} \approx 16\pi^2\ell_P \sum_{n\geq 0} \frac{a_n^2 + b_n^2}{(n+1)(n+2)(n+3)} + \frac{\ell\ell_{AdS}}{8\pi\ell_P\sqrt{2\chi}} \tag{8.67}$$

That is, in this regime, the system behaves as a (countably infinite) collection of independent free "Newtonian particles", each with kinetic energy $p_n^2/2M_n$, where the "effective mass" is given by

$$M_n \sim \frac{(n+1)(n+2)(n+3)}{32\pi^2\ell_P} \tag{8.68}$$

However, we should only read off the "effective masses" after expressing the coefficients a_n and b_n in terms of the canonical momenta P_n, which is a topic for Part II. This detail aside, the masses above tend to grow with the mode number n, which is satisfying since this means that the modes associated with a more oscillatory diamond corner[37] are more massive, and thus more difficult to "excite". Finally, we remark that this approximate form of the Hamiltonian should be interpreted with care, as we must not forget that the exact Hamiltonian, although bounded from

[37] As we are going to discuss later, the "momentum variables" are related to how the corner loop of the diamond wiggles in a lightlike direction, while the "position variables" are related to how it wiggles in a spacelike direction. This interpretation also provides a way to see that "larger p" tends to increase the Hamiltonian (since a more lightlike boundary can enclose a larger area for a given perimeter) and a "larger q" tends to decrease the Hamiltonian (since more spatial wiggling tends to shrink the enclosed area).

below, does not have a minimum (similar to an exponential, $e^x > 0$). Therefore, the symmetric diamond is not the "ground state", as one could perhaps think by just looking at this approximated Hamiltonian, but there are states with lower "energy".

It would be interesting to carry out the Hamiltonian expansion up to higher orders, where higher terms in $e^{\delta\lambda}$ would be kept. At second order we expect a non-trivial potential $V(\psi)$ to appear. Note that, at first order, one could be erroneously led to believe that one is expanding around a minimum of the Hamiltonian, since $p^2/2M$ is minimized at $p = 0$. However, we know, non-perturbatively, that the symmetric state $(\psi, \mathring{\sigma}) = (I, 0)$ is not a minimum of the Hamiltonian, but rather a saddle point. In fact, since the Hamiltonian corresponds to the area of the CMC slice, it is bounded from below (by zero) but it has no minimum (since there is no regular state with exactly zero area). This feature is reminiscent of Liouville field theory, whose potential has an exponential form, $V(\phi) = e^{2b\phi}$, and has shown many links to low-dimensional gravity [73, 74, 75, 76, 77, 78, 79].

Part II

Group-theoretic quantization

9 Group-theoretic quantization

Canonical quantization is a fantastical tool discovered by Dirac that allows one to infer the quantum theory underlying a given classical theory. It is based on the idea that the dynamical laws of classical and quantum theories, and the underlying mathematical structures describing those laws, are fundamentally analogous. The traditional approach to quantization usually involves a choice of conjugate coordinates, x's and p's, on the phase space which are promoted to quantum operators satisfying the canonical commutation relations. However, this approach does not apply directly to phase spaces with nontrivial topologies, or lacking a natural linear structure, since they do not admit a (natural) global coordinate chart. Since the reduced phase space of the diamond, $T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$, does not have a natural linear structure, we must resort to more general canonical quantization schemes. In particular, as this phase space is the cotangent bundle of a homogeneous space, it is natural to consider Isham's group-theoretic approach to quantization [41, 42] in which the quantum theory is based on a transitive group of symmetries of the phase space. This section contains a brief introduction to the basic principle of canonical quantization, followed by a general review of Isham's group-theoretic quantization scheme, which is finally specialized to the case of phase spaces with a cotangent bundle structure.

9.1 The basic canons of quantization

Canonical quantization posits that the quantum theory should retain the dynamical laws of the classical theory, to the highest possible extent, while replacing the classical notion of kinematics by the appropriate quantum notion, where states go from points in a phase space to vectors (or rather, rays) in a Hilbert space and observables go from real functions on the phase space to self-adjoint operators on the Hilbert space. More precisely, classical observables are real functions f on a phase space, endowed with a Poisson algebra coming from the symplectic structure,

$$f *_C g := \{f, g\} \tag{9.1}$$

in a corresponding quantum theory, observables are self-adjoint operators $\widehat{f}$ on a Hilbert space, endowed with a commutator algebra

$$\widehat{f} *_Q \widehat{g} := \frac{1}{i\hbar}[\widehat{f}, \widehat{g}] \tag{9.2}$$

where the complex unit is included so that the product returns a self-adjoint operator and $\hbar$ is a dimensional constant (interpreted as the Planck constant). The principle of canonical quantization is that given an appropriate subalgebra of observables in the classical theory, $\mathcal{A}_C$, there would be an homomorphism $\mathfrak{q}$ to a corresponding subalgebra of operators in the quantum theory, $\mathcal{A}_Q$. That is, $\mathfrak{q} : \mathcal{A}_C \to \mathcal{A}_Q$ is a linear association from classical observables to quantum observables, $f \mapsto \widehat{f} := \mathfrak{q}(f)$, such that $\mathfrak{q}(f *_C g) = \mathfrak{q}(f) *_Q \mathfrak{q}(g)$, or, in the more familiar notation, $[\widehat{f}, \widehat{g}] = i\hbar\,\widehat{\{f, g\}}$. The intention is to preserve, to a sensible extent, the dynamical structure of the theory. Namely, the time evolution of a classical observable f under the Hamiltonian H is given by

$$\frac{df}{dt} = \{f, H\} \tag{9.3}$$

while, in the quantum theory, the time evolution (in the Heisenberg picture) of an observable $\widehat{f}$ under the Hamiltonian $\widehat{H}$ is given by

$$\frac{d\widehat{f}}{dt} = \frac{1}{i\hbar}[\widehat{f}, \widehat{H}] \tag{9.4}$$

indicating that mapping $\{f, H\} \mapsto \frac{1}{i\hbar}[\widehat{f}, \widehat{H}]$ is precisely what is needed to ensure that df/dt is mapped to $d\widehat{f}/dt$, so that the *quantization map* is preserved in time. However, as explained below, there are often obstructions in constructing such a quantization map, consistently, for all observables. Accordingly, different quantization schemes propose different manners to construct a "most consistent" quantization map.

In general, the classical subalgebra of observables $\mathcal{A}_C$ is assumed to be complete in the sense that the specification of the value of all $f \in \mathcal{A}_C$ completely determines the state (i.e., point in the phase space), possibly up to a finite ambiguity (i.e., any point in the phase space has a neighborhood whose points are uniquely determined by the values of the f's). Such a subalgebra of observables, when selected as the basis for quantization, will be called a *canonical algebra*, and its elements *canonical observables* or also *canonical charges*. It would be desirable if this homomorphism could be taken to satisfy some basic properties, such as:

(i) Mapping the constant unit function, $u : \mathcal{P} \to \{1\} \subset \mathbb{R}$, to the identity operator, i.e., $\mathfrak{q}(u) = 1$;

(ii) Commutativity with composition by real functions, i.e., for any real function $\phi : \mathbb{R} \to \mathbb{R}$, $\mathfrak{q}(\phi(f)) = \phi(\mathfrak{q}(f))$. (A.k.a., the von Neumann rule.)

One may consider these properties desirable for the following reasons. First, an observable that takes the same value 1 at every classical state should correspond to an operator whose spectrum contains the single value 1, which is the identity. Second, if one has designed a classical experiment to measure a given observable f, then the same experiment can also be used to measure any other function ϕ of f by simply running it and acting on the output data with ϕ; at the quantum level one could expect that the same principle is true, in that if this experiment is run and the system is found to be in an eigenvector $|\alpha\rangle$ of $\mathfrak{q}(f)$ with eigenvalue α, then the same experiment could be used to measure $\mathfrak{q}(\phi(f))$, and the same state $|\alpha\rangle$ would also be an eigenvector of the corresponding quantum operator with eigenvalue $\phi(\alpha)$ — but, by definition, the operator that has the same eigenvectors $|\alpha\rangle$ as $\mathfrak{q}(f)$, with corresponding eigenvalues $\phi(\alpha)$, is precisely $\phi(\mathfrak{q}(f))$. In principle, this second property would allow one to extend the quantization map to the entire algebra of observables, since any observable can be (locally) written in terms of a complete set of observables — such an extension is not unique due to operator-ordering ambiguities. In addition, the corresponding notion of completeness at the quantum level is that the algebra must be represented irreducibly in the Hilbert space. This is based on the correspondence between symplectomorphisms, at the classical level, with unitary transformations, at the quantum level. Namely, an algebra of classical observables is complete, in a sense equivalent to the discussed above, if the symplectomorphisms it generates act transitively on the phase space, i.e., any two states can be connected by "evolving along" some observable; at the quantum level, an algebra is represented irreducibly if there is no (non-trivial) invariant subspace, which implies that by acting on a state with all elements of the algebra (and their products) the whole Hilbert space is spanned, i.e., any[38] two quantum states can be connected by the unitary exponentiation of some observable in the algebra.

Naturally one would first attempt quantizing the entire algebra of observables, but fundamental obstructions are generally encountered, as revealed by Groenewold-van Hove no-go theorem.[39] In general, there is no quantization map that commutes with the composition by real functions (i.e., that satisfies the von Neumann rule).

[38]Rigorously, the orbit of a state may not be the full Hilbert space, but rather a dense subset.

[39]Strictly speaking, Groenewold-van Hove theorem, in its standard formulation, applies only for trivial phase spaces, $\mathcal{P} = \mathbb{R}^{2n}$. See [80] for details. The result has been extended to other cases, and it is expected that this kind of obstruction is generic [81]. However, some examples have been found where a full, unobstructed quantization is possible [82, 83].

Moreover, one would expect that a quantization is sensible only if it preserves the natural relationship between the classical and the quantum notions of algebra completeness discussed above. That is, if $\mathcal{A}_C$ is a complete algebra of classical observables and there exists a subalgebra $\mathcal{A}'_C \subset \mathcal{A}_C$ that is also complete at the classical level, then the quantization $\mathcal{A}_C \mapsto \mathcal{A}_Q$, where $\mathcal{A}_Q$ is represented unitarily and irreducibly on a Hilbert space, naturally induces a unitary representation of $\mathcal{A}'_Q$ on the same Hilbert space, by restricting to the corresponding subalgebra of $\mathcal{A}_Q$, and it would be physically reasonable if this representation were also irreducible. Nonetheless, it is not true in general that there exists an irreducible representation of $\mathcal{A}_Q$ such that all of its classically-complete subalgebras are represented irreducibly (and therefore also quantum-complete).

The traditional proposal by Dirac applies well for linear phase spaces, i.e., those isomorphic to a vector space $\mathbb{R}^{2n}$, with some preferred set of global conjugate coordinates x^i and p_i, satisfying $\{x^i, p_j\} = \delta^i{}_j$. These $2n$ coordinates, together with the constant function 1, form a Heisenberg algebra of observables that is complete. Therefore, they could be used as basis for quantization, that is, the quantum theory would be based on a unitary irreducible representation of the Heisenberg algebra, $[X^i, P_j] = i\hbar\delta^i{}_j$. There are two important aspects of this quantization that must be stressed. First, the quantum theory is only sensible if the coordinates are global and range from $-\infty$ to ∞. The reason is that, according to Stone-von Neumann theorem [84], the Heisenberg algebra has a unique (up to unitary equivalence) irreducible unitary representation: the states are described by the familiar $\mathbb{C}$-valued, square-integrable wave-functions on $\mathbb{R}^n$, $\Psi(x)$, on which the canonical observables act as

$$
\begin{aligned}
(X^i\Psi)(x) &= x^i\Psi(x) \\
(P_i\Psi)(x) &= -i\hbar\frac{\partial\Psi}{\partial x^i}(x)
\end{aligned}
\tag{9.5}
$$

The spectrum of X^i is therefore the whole real line.[40] Second, the quantization is only natural insofar as the coordinates are natural, for generally different choices of coordinates may lead to inequivalent quantum theories (e.g., because of operator ordering issues associated with non-linear transformations between coordinates). In the case of a particle on an Euclidean plane, one may argue that Cartesian spatial coordinates, selected by the flat metric, paired with their conjugate momenta, provide a natural basis for quantization. (In fact, different Cartesian systems are linearly related, so all such quantizations are equivalent.)

If the phase space is non-trivial, in the sense that there is no natural global coordinate system to employ, then one must be more careful to carry out the quantization. Two very simple examples are: a particle living on a half-line or on a

[40] Another simple manner to see this is the following. Suppose there are self-adjoint operators X and P satisfying the algebra $[X,P]=i$. Since P is self-adjoint, it can be exponentiated, so $e^{-iaP/\hbar}$ is a well-defined bounded operator for any $a\in\mathbb{R}$. Now suppose that X has a eigenvector $|x\rangle$, with eigenvalue $x\in\mathbb{R}$. Then it derives from the algebra that $e^{-iaP/\hbar}|x\rangle$ is an eigenvector of X with eigenvalue $x+a$. Therefore, if the spectrum of X is not empty (so the representation is non-trivial), then it must be the whole $\mathbb{R}$. Rigorously, X does not have eigenvectors in the strict sense, since "would-be" eigenvectors associated with a continuum spectrum are non-normalizable, but it still has eigenvectors in a limiting sense, i.e., if λ is in the spectrum then there exists a sequence of non-zero vectors Ψ_n such that $\lim_{n\to\infty}\|(X-\lambda)\Psi_n\|/\|\Psi_n\|=0$ [84].

sphere, where the phase spaces are respectively $T^*\mathbb{R}^+$ and T^*S^2. In the first case, it may appear that one could take coordinates $x > 0$ and p, still satisfying $\{x, p\} = 1$, but as we have seen the quantization would not be sensible since there is no unitary representation of this algebra in which the spectrum of X is restricted to positive numbers. In the second case, the non-trivial topology of the sphere implies that there is no global system of coordinates. When there is no global chart, one could still try to quantize a local chart, which may be sensible in certain approximations, but global and topological aspects would be missed. In particular, the spin variable $S = J \cdot N$, where J is the angular momentum and N is the unit vector describing the (angular) position of the particle, turns out to be quantized in integer multiples of $\hbar/2$, dissimilar to what happens for a particle on a plane $\mathbb{R}^2$ (i.e., local chart) where the spin is continuous — in fact, this quantization of spin is directly related to Dirac's quantization of the electric-magnetic monopoles, $eg = n\hbar/2$. It is worthwhile to note that the Stone-von Neumann theorem, establishing uniqueness of the representation, is very particular to the Heisenberg algebra, which in turn is very particular to trivial phase spaces — in other cases, the kinematical part of the quantization (i.e., finding a self-adjoint representation of canonical variables) is itself a potentially complicated problem, and the structure of the Hilbert space is only determined after this stage is resolved (for example, in the case of a particle on the sphere, different spin values correspond to inequivalent quantum representations [47]).

In view of these complications, different frameworks have been developed to employ the canonical principles of quantization to nontrivial phase spaces. Among them we can mention geometric quantization [85, 86, 87, 88, 89], deformation quantization [90, 91, 91] and group-theoretic quantization [41, 42, 92, 93]. They all offer a prescription that attempts to find a reasonable balance that minimizes the extent that the classical and quantum algebras fail to be homomorphic, that the von Neumann rule is violated, and that the classical and quantum notions of completeness disagree. Due to the form of the reduced phase space for causal diamonds, revealing a natural underlying group structure, we shall focus on Isham's group-theoretic formalism.

9.2 Isham's quantization scheme

Isham proposed that one should identify a transitive group of symplectic symmetries of the phase space and use it to generate both a special set of classical observables and their associated quantum self-adjoint operators [41, 42]. Such a group will be called the *quantizing group*, or also the *canonical group*. For intuition, see the introduction (Sec. 1.2.2) where we discuss the trivial case of a particle on the line, i.e., a phase space $T^*\mathbb{R} = \mathbb{R}^2$. Let us now review the general formalism underlying this approach.

Consider a phase space $\widetilde{\mathcal{P}}$ with symplectic 2-form ω. Assume that the phase space is a homogeneous space for some Lie group $\widetilde{G}$ of symplectic symmetries.[41] That is, there is a left action of $\widetilde{G}$ on $\widetilde{\mathcal{P}}$, $\Gamma_{\widetilde{g}} : \widetilde{\mathcal{P}} \to \widetilde{\mathcal{P}}$, that preserves the symplectic form,

$$\Gamma^*_{\widetilde{g}}\omega = \omega \tag{9.6}$$

for all $\widetilde{g} \in \widetilde{G}$, and is transitive, i.e., given any two points $p, p' \in \widetilde{\mathcal{P}}$ there exists $\widetilde{g} \in \widetilde{G}$ such that $p' = \Gamma_{\widetilde{g}}(p)$. Each element $\widetilde{\xi}$ in the Lie algebra $\widetilde{\mathfrak{g}}$ of $\widetilde{G}$ induces a vector field $X_{\widetilde{\xi}}$ on $\widetilde{\mathcal{P}}$, as follows. The algebra $\widetilde{\mathfrak{g}}$ is naturally identified with the tangent space of $\widetilde{G}$ at the identity, $T_e\widetilde{G}$, so $\widetilde{\xi}$ can be seen as a vector tangent to a curve $\widetilde{g}_t$ on $\widetilde{G}$ starting at e; acting with this curve on any point $p \in \widetilde{\mathcal{P}}$ defines a curve $p_t := \Gamma_{\widetilde{g}_t}(p)$ on $\widetilde{\mathcal{P}}$ starting at p, and therefore a vector $X_{\widetilde{\xi}}$ tangent to p. More formally,

$$X_{\widetilde{\xi}}|_p := \phi_{p*}(\widetilde{\xi}) \tag{9.7}$$

where $\phi_p : \widetilde{G} \to \widetilde{\mathcal{P}}$ is defined by $\phi_p(g) = \Gamma_{\widetilde{g}}(p)$. This map is an anti-homomorphism from $\widetilde{\mathfrak{g}}$ into the algebra of vector fields on $\widetilde{\mathcal{P}}$, i.e.,

$$[X_{\widetilde{\xi}}, X_{\widetilde{\eta}}] = X_{[\widetilde{\eta},\widetilde{\xi}]} \tag{9.8}$$

As $\Gamma_{\widetilde{g}}$ preserves ω, we have $\pounds_{X_{\widetilde{\xi}}}\omega = 0$, so $X_{\widetilde{\xi}}$ is a (locally) Hamiltonian field. Thus $d(\imath_{X_{\widetilde{\xi}}}\omega) = \pounds_{X_{\widetilde{\xi}}}\omega - \imath_{X_{\widetilde{\xi}}}d\omega = 0$, where $\imath$ denotes the interior product. That is, $\imath_{X_{\widetilde{\xi}}}\omega$

[41]The tilde will typically indicate symbols referring to the phase space $\widetilde{\mathcal{P}}$, such as the group $\widetilde{G}$. Symbols referring to the configuration space $\mathcal{Q}$, such as the group of "translations" G in the next subsection, will typically be denoted without accents. While the tildes cause an unfortunate clutter in this section, this choice of notation will prove convenient in the rest of the paper.

is closed and therefore locally exact, so that

$$dH_{\widetilde{\xi}} = -\imath_{X_{\widetilde{\xi}}}\omega \tag{9.9}$$

admits local solutions $H_{\widetilde{\xi}}$, called *canonical charges*, defined up to addition of a constant function on $\widetilde{\mathcal{P}}$. In order to properly regard these charges as canonical observables, they need to be globally defined on $\widetilde{\mathcal{P}}$, so we assumed that $\widetilde{G}$ strictly generates *globally* Hamiltonian fields on $\widetilde{\mathcal{P}}$. This set of charges will be (classically) complete as a consequence of the transitiveness of the group action.

The symplectic form endows the space of functions on the phase space with an algebraic structure, $\mathcal{A}_C$, where the product is given by the Poisson bracket: since ω is non-degenerate, any function f on $\widetilde{\mathcal{P}}$ can be associated with a unique vector field X_f on $\widetilde{\mathcal{P}}$ via the relation $df = -\imath_{X_f}\omega$; the Poisson bracket between two functions, f and f', is defined by

$$\{f, f'\} := -\omega(X_f, X_{f'}) \tag{9.10}$$

It is straightforward to show that the vector field associated with $\{f, f'\}$ is $-[X_f, X_{f'}]$, so there is an anti-homomorphism from the Poisson algebra of charges to the algebra of their associated vector fields. For the canonical charges, $\{H_{\widetilde{\xi}}, H_{\widetilde{\eta}}\}$ is then associated with the vector field $-[X_{\widetilde{\xi}}, X_{\widetilde{\eta}}] = X_{[\widetilde{\xi},\widetilde{\eta}]}$, which thus imply that $d(\{H_{\widetilde{\xi}}, H_{\widetilde{\eta}}\} - H_{[\widetilde{\xi},\widetilde{\eta}]}) = 0$. Therefore the map $\widetilde{\xi} \mapsto H_{\widetilde{\xi}}$ is a homomorphism from $\widetilde{\mathfrak{g}}$

into $\mathcal{A}_C$ up to central charges, i.e.,

$$\{H_{\widetilde{\xi}}, H_{\widetilde{\eta}}\} = -\omega(X_{\widetilde{\xi}}, X_{\widetilde{\eta}}) = H_{[\widetilde{\xi},\widetilde{\eta}]} + z(\widetilde{\xi},\widetilde{\eta}) \tag{9.11}$$

where $z(\widetilde{\xi},\widetilde{\eta})$ is constant on $\widetilde{\mathcal{P}}$. If the central charge $z(\widetilde{\xi},\widetilde{\eta})$ is not trivial (i.e., it cannot be removed by a redefinition of the charges $H_{\widetilde{\xi}} \mapsto H_{\widetilde{\xi}} + f(\widetilde{\xi})$, for some $f : \widetilde{\mathfrak{g}} \to \mathbb{R}$), we can always extend the group by a central element so that the extended algebra, $\widehat{\widetilde{\mathfrak{g}}}$, with topology $\widetilde{\mathfrak{g}} \oplus \mathbb{R}$, has product law $[(\widetilde{\xi}; a), (\widetilde{\eta}; b)] = ([\widetilde{\xi},\widetilde{\eta}]; z(\widetilde{\xi},\widetilde{\eta}))$. The new group (obtained by exponentiating $\widehat{\widetilde{\mathfrak{g}}}$) has a natural action on $\widetilde{\mathcal{P}}$, where the central element acts trivially, and the new charges are related to the old ones simply by $H_{(\widetilde{\xi};a)} = H_{\widetilde{\xi}} + a$. Consequently, the map $(\widetilde{\xi}, a) \mapsto H_{(\widetilde{\xi},a)}$ is a true homomorphism. In this way, if we start with a candidate canonical group that generates non-trivial central charges when acting on the phase space, then we can restart the process with the appropriately extended group as the canonical group. Thus, we will assume $\widetilde{G}$ is chosen such that

$$\{H_{\widetilde{\xi}}, H_{\widetilde{\eta}}\} = H_{[\widetilde{\xi},\widetilde{\eta}]} \tag{9.12}$$

i.e., so that $\widetilde{\mathfrak{g}} \to \mathcal{A}_C$ is a true homomorphism.[42]

[42]For example, in the case of $\widetilde{\mathcal{P}} = T^*\mathbb{R}$, the natural group to consider is $\mathbb{R}^2$ acting as translations on $\widetilde{\mathcal{P}}$, i.e., $(a, b) \in \widetilde{G}$ acting on $(x, p) \in \widetilde{\mathcal{P}}$ as $\Gamma_{(a,b)}(x, p) = (x + a, p + b)$. Its algebra $\widetilde{\mathfrak{g}} = \mathbb{R}^2$ is commutative, and the associated charges can be taken as $H_{(\alpha,\beta)} = \alpha p - \beta x$, where $(\alpha, \beta) \in \widetilde{\mathfrak{g}}$. A non-trivial central charge appears in the Poisson algebra, $\{H_{(\alpha,\beta)}, H_{(\alpha',\beta')}\} = \alpha\beta' - \beta\alpha'$. The algebra can be extended by the central element to $\widehat{\widetilde{\mathfrak{g}}}$ with product law $[(\alpha, \beta; \gamma), (\alpha', \beta'; \gamma')] := (0; \alpha\beta' - \beta\alpha')$, known as the (3-dimensional) Heisenberg algebra, $\mathfrak{h}(3)$. The exponentiation of this algebra defines the Heisenberg group, $H(3) = \widehat{\mathbb{R}^2}$, with topology $\mathbb{R}^3$, whose product rule can be expressed as $(a, b; c)(a', b'; c') = (a + a', b + b'; c + c' + \frac{1}{2}ab' - \frac{1}{2}ba')$. Thus the quantization should be based on $\widetilde{G} = H(3)$, instead of the original candidate $\mathbb{R}^2$. This group acts on $\widetilde{\mathcal{P}}$ as $\Gamma_{(a,b;c)}(x, p) = (x+a, p+b)$ and the associated charges are taken to be $H_{(\alpha,\beta;\gamma)} = \alpha p - \beta x + \gamma$.

The quantum theory is constructed from $\widetilde{G}$ as follows. Let $U : \widetilde{G} \to \mathrm{Aut}(\mathcal{H})$ be an irreducible unitary representation of $\widetilde{G}$ on a Hilbert space $\mathcal{H}$. As any algebra element $\widetilde{\xi} \in \widetilde{\mathfrak{g}}$ generates a one-parameter subgroup of $\widetilde{G}$ through exponentiation, $t \mapsto \exp(t\widetilde{\xi})$, they can be associated with self-adjoint operators $\widehat{H}_{\widetilde{\xi}}$ on $\mathcal{H}$ via

$$U(\exp t\widetilde{\xi}) =: e^{t\widehat{H}_{\widetilde{\xi}}/i\hbar} \tag{9.13}$$

It follows from this definition that the map $\widetilde{\xi} \mapsto \widehat{H}_{\widetilde{\xi}}$ is a homomorphism from $\widetilde{\mathfrak{g}}$ into $\mathcal{A}_Q$, the algebra of self-adjoint operators on $\mathcal{H}$,

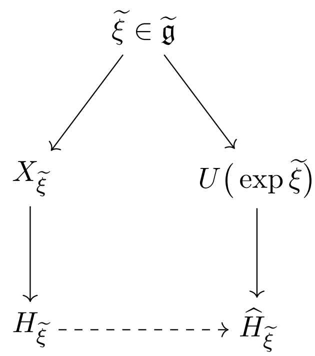
$$\frac{1}{i\hbar}[\widehat{H}_{\widetilde{\xi}}, \widehat{H}_{\widetilde{\eta}}] = \widehat{H}_{[\widetilde{\xi},\widetilde{\eta}]} \tag{9.14}$$

Therefore, considering (9.12) and (9.14), we see that the the association between each classical charge $H_{\widetilde{\xi}}$ and the corresponding generator $\widehat{H}_{\widetilde{\xi}}$ of the unitary representation,

$$H_{\widetilde{\xi}} \mapsto \widehat{H}_{\widetilde{\xi}}, \tag{9.15}$$

is a homomorphism from $\mathcal{A}_C$ into $\mathcal{A}_Q$. This is called the *quantization map*, for the quantization based on the group $\widetilde{G}$.

The quantization process is summarized as follows:



A transitive group $\widetilde{G}$ of symplectomorphisms defines classical charges and their corresponding quantum operators. On the classical side, each element $\widetilde{\xi}$ in the Lie algebra $\widetilde{\mathfrak{g}}$ induces a Hamiltonian vector field $X_{\widetilde{\xi}}$ on $\widetilde{\mathcal{P}}$, which in turn defines a Hamiltonian charge $H_{\widetilde{\xi}}$. On the quantum side, the group element $\exp \widetilde{\xi}$ is represented by a unitary operator on $\mathcal{H}$, whose self-adjoint generator is $\widehat{H}_{\widetilde{\xi}}$. The association $H_{\widetilde{\xi}} \mapsto \widehat{H}_{\widetilde{\xi}}$ is, by construction, a homomorphism between $\mathcal{A}_C$ and $\mathcal{A}_Q$.

As global phases are unphysical, the true space of quantum states is the *ray space*, $\mathcal{R} := \mathcal{H}/U(1)$, corresponding to the quotient of the Hilbert space $\mathcal{H}$ by phases $e^{i\theta} \in U(1)$. Accordingly, symmetries must be represented by unitary operators up to a phase. These are known as *projective representations* of the group $\widetilde{G}$, more precisely defined as homomorphisms from $\widetilde{G}$ to the group of projective unitary operators on $\mathcal{R}$, $P\mathscr{U}(\mathcal{H}) := \{U \sim e^{i\theta}U;$ where $U \in \mathscr{U}(\mathcal{H})$ and $\theta \in \mathbb{R}\}$. It is a general theorem (reviewed in App. F) that projective irreducible unitary representations of $\widetilde{G}$ are in one-to-one correspondence irreducible self-adjoint representations of a central extension of the Lie algebra $\widetilde{\mathfrak{g}}$ by 2-cocycles. As explained later, we will append to the quantization prescription a principle of Casimir matching, which implies that if $\widetilde{\mathfrak{g}}$ is properly represented on the phase space as in (9.12), without additional central charges, then the quantum theory should be constructed from irreducible representations of $\widetilde{\mathfrak{g}}$, without further extensions. Equivalently, the quantum theory should be constructed from unitary irreducible representations of the universal cover of $\widetilde{G}$. In this manner, we can understand Isham's method as a prescription to generate an algebraically-closed complete set of classical observables, $H_{\widetilde{\xi}}$, associated to some symmetry structure of the phase space, which is then quantized according to the standard rule of canonical quantization, where the same algebra is irreducibly represented by corresponding self-adjoint operators $\widehat{H}_{\widetilde{\xi}}$ on a Hilbert space.

We mentioned that the reason for requiring a *transitive* action on the phase space is so that the set of generated charges is classically complete, i.e., that any function on $\widetilde{\mathcal{P}}$ can be (locally) expressed in terms of the $H_{\widetilde{\xi}}$'s. In fact, note that transitivity implies that, at any $p \in \widetilde{\mathcal{P}}$, any tangent vector $X \in T_p\widetilde{\mathcal{P}}$ is equal to $X_{\widetilde{\xi}}$ for some $\widetilde{\xi} \in \widetilde{\mathfrak{g}}$. Then, for any $V \in T_p\widetilde{\mathcal{P}}$, $dH_{\widetilde{\xi}}(V) = -\omega(X_{\widetilde{\xi}}, V)$ will not vanish for at least one $\widetilde{\xi}$, since ω is non-degenerate. In words, there is no direction V along which all charges are (locally) constant, confirming that any function on $\widetilde{\mathcal{P}}$ can be locally written in terms of the charges. A more geometrical way to see this involves the momentum map, $\mu : \widetilde{\mathcal{P}} \to \widetilde{\mathfrak{g}}^*$, defined by $\mu(p)(\widetilde{\xi}) := H_{\widetilde{\xi}}(p)$, where $p \in \widetilde{\mathcal{P}}$ and $\widetilde{\xi} \in \widetilde{\mathfrak{g}}$. The observation above implies that μ_* is injective[43] and therefore μ is an immersion. Consequently, any $p \in \widetilde{\mathcal{P}}$ has a neighborhood $\mathcal{U} \subset \widetilde{\mathcal{P}}$ such that $\mu|_{\mathcal{U}}$ is

[43]Since $\widetilde{\mathfrak{g}}^*$ is a vector space, tangent vectors can be identified with elements of $\widetilde{\mathfrak{g}}^*$ itself. In this way, $\mu_* : T_p\widetilde{\mathcal{P}} \to \widetilde{\mathfrak{g}}^*$ is given by $\mu_*(V)(\widetilde{\xi}) = dH_{\widetilde{\xi}}(V) = -\omega(X_{\widetilde{\xi}}, V)$, so $\mu_*(V)$ is zero if and only if $V = 0$.

an embedding of $\mathcal{U}$ into $\widetilde{\mathfrak{g}}^*$. Now, given any basis $\widetilde{\xi}_i$ for $\widetilde{\mathfrak{g}}$, there is an associated coordinate system $\{\widetilde{\alpha}_i\}$ on $\widetilde{\mathfrak{g}}^*$ defined by $\widetilde{\alpha}_i|_{\widetilde{\alpha}} := \widetilde{\alpha}(\widetilde{\xi}_i)$, where $\widetilde{\alpha} \in \widetilde{\mathfrak{g}}^*$. It follows that $\mu^*\widetilde{\alpha}_i = H_{\widetilde{\xi}_i}$. Given any smooth real function $f : \mathcal{U} \to \mathbb{R}$, there exists a neighborhood $\overline{\mathcal{U}}$ of $\mu(\mathcal{U})$ and a smooth real function $\overline{f} : \overline{\mathcal{U}} \to \mathbb{R}$, such that $f = \mu^*\overline{f}$. As $\overline{f}$ can be written in terms of the coordinates $\widetilde{\alpha}_i$, f can be written in terms of the charges $H_{\widetilde{\xi}_i}$.[44]

The prescription, as formulated by Isham, is intended to provide a general framework for quantization and, as such, refers only to the minimal structure necessary for a "kinematical quantization", i.e., producing sensible homomorphism between a classically-complete subalgebra of observables to a corresponding quantum-complete subalgebra of self-adjoint operators. The job is not finished, however, until a particular representation of the canonical observables is determined and other physically relevant observables (which are classically expressed in terms of the canonical ones) are also promoted to self-adjoint operators on the Hilbert space. In particular, the time-evolution Hamiltonian, describing the dynamics of the system, may often not be part of the canonical algebra, and need to be included in the quantum theory eventually. These additional observables that may need to be quantized often help selecting a preferred class of irreducible representations of the canonical algebra, therefore partially reducing the ambiguities in the quantization.

Another rule that we employ for filtering the representations is the *Casimir matching principle*. Casimir operators are elements in the center of a universal

[44] If $\overline{f}(\widetilde{\alpha}) = F(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots)$, for some $F : \mathbb{R}^{dim(\widetilde{\mathfrak{g}})} \to \mathbb{R}$, then $f(p) = F(H_{\widetilde{\xi}_1}(p), H_{\widetilde{\xi}_2}(p), \ldots)$.

enveloping algebra.[45] At the classical level, the completeness of the algebra implies that Casimir observables are constant on the phase space. At the quantum level, Schur's lemma implies that in any irreducible (complex) representation Casimir operators are multiples of the identity. At both levels, Casimirs take the same value for all states. Thus, if the classical system is to arise from a classical limit of the quantum system, the eigenvalue of the quantum Casimir observable should match the value of the corresponding classical Casimir. In the study of the quantization of a particle on a sphere, in the presence of a magnetic monopole, from the perspective of Isham's method, this principle plays a key role in the argument leading to Dirac's charge quantization condition, $eg = n\hbar/2$ [47].

9.3 Phase spaces with a cotangent bundle structure

Now we discuss a general method, also described by Isham [41, 42], for constructing a transitive group of symplectomorphisms of phase spaces that are the cotangent bundle of homogeneous manifolds. Consider a phase space $\widetilde{\mathcal{P}} = T^*\mathcal{Q}$, with the canonical symplectic form $\omega = d\theta$ associated with the cotangent bundle structure, where the configuration space $\mathcal{Q}$ is homogeneous space for some Lie group G. Let

[45] A universal enveloping algebra is a formal extension of a Lie algebra consisting of polynomials of the algebra elements, where the product is defined by treating the elements as operators in an abstract representation. For example, say that α, β, γ and δ are elements of $\mathfrak{g}$, then $\alpha\beta + \gamma$ belongs to the universal enveloping algebra of $\mathfrak{g}$ and the product with δ is given by $[\alpha\beta + \gamma, \delta] := \alpha[\beta, \delta] + [\alpha, \delta]\beta + [\gamma, \delta]$.

us denote the left G-action on $\mathcal{Q}$ by $\delta : G \to \mathit{Diff}(\mathcal{Q})$, or simply by

$$gx := \delta_g(x) \tag{9.16}$$

where $x \in \mathcal{Q}$. There is a natural lift of this action to the cotangent bundle, $\widetilde{\delta} : G \to \mathit{Diff}(\widetilde{\mathcal{P}})$, given by

$$\widetilde{\delta}_g(p) := \delta^*_{g^{-1}} p \tag{9.17}$$

which maps the fiber over $x = \pi(p)$, where $\pi : \widetilde{\mathcal{P}} \to \mathcal{Q}$ is the bundle projection, to the fiber over gx. That is, the lifted action satisfies the identity $\pi \circ \widetilde{\delta}_g = \delta_g \circ \pi$. The symplectic potential θ is invariant under such a transformation, which can be seen as follows. If $V \in T_p\widetilde{\mathcal{P}}$, then $\widetilde{\delta}^*_g\theta(V) = \theta(\widetilde{\delta}_{g*}V)$, and from the definition of θ this is $(\widetilde{\delta}_g(p))(\pi_*\widetilde{\delta}_{g*}V)$; from the definition of the lifted action this is equal to $\delta^*_{g^{-1}}p(\pi_*\widetilde{\delta}_{g*}V) = p(\delta_{g^{-1}*}\pi_*\widetilde{\delta}_{g*}V)$, and using the identity above we get $p(\delta_{g^{-1}*}\delta_{g*}\pi_*V) = p(\pi_*V) = \theta(V)$, that is, $\widetilde{\delta}^*_g\theta = \theta$. Consequently, $\widetilde{\delta}^*_g d\theta = d\widetilde{\delta}^*_g\theta = d\theta$, i.e.,

$$\widetilde{\delta}^*_g\omega = \omega \tag{9.18}$$

so the group G, associated with $\mathcal{Q}$, acts as symmetries of the phase space.

The group G alone cannot serve as the quantizing group though, since it does not act transitively on the phase space. Rather, it acts only "laterally" on the phase space, so we must enlarge the group by including also transformations along the fibers of the bundle. A natural choice for a "vertical" transformation is to consider *momentum translations* by 1-forms, that is, given a 1-form field α on $\mathcal{Q}$, define

$$\zeta_\alpha(p) := p - \alpha \tag{9.19}$$

where α on the right-hand side is evaluated at $\pi(p)$. We can compute how the symplectic potential changes under this transformation as follows. Given any $V \in T_p\widetilde{\mathcal{P}}$, we have

$$\zeta_\alpha^*\theta(V) = \theta(\zeta_{\alpha *}V) = (\zeta_\alpha p)(\pi_*\zeta_{\alpha *}V) = (p - \alpha)(\pi_* V) = (\theta - \pi^*\alpha)(V) \tag{9.20}$$

where we have used that ζ_α acts vertically and thus satisfies the identity $\pi \circ \zeta_\alpha = \pi$. The symplectic 2-form then transforms like

$$\zeta_\alpha^*\omega = \omega - \pi^* d\alpha \tag{9.21}$$

Therefore, in order to have a symmetry we must restrict to closed 1-form fields, $d\alpha = 0$. As we shall see, it is necessary to restrict further to exact 1-forms, $\alpha = df$, where $f \in C^\infty(\mathcal{Q}, \mathbb{R})$, so that the associated charges are globally defined on the phase space.

The space of all exact 1-form fields on $\mathcal{Q}$, which is infinite-dimensional, is unnecessarily large for our purposes, as we are interested in constructing a "minimal" transitive group of symmetries acting on the phase space. That is, given the principle discussed in Sec. 9.1 that the classical and quantum notions of algebra completeness must agree, we should always focus on canonically quantizing an algebra that has no proper classically-complete subalgebras, so that this principle is vacuously satisfied. (Note the analogy with the case of a trivial phase space, where only the Heisenberg algebra is canonically quantized, as opposed to the whole algebra of classical observables.) If $\mathcal{Q}$ is n-dimensional, we would hope that "just about" n generators would already suffice to produce a transitive action (of course, we need at least n). There is a natural algorithm to select an appropriate subset of exact 1-forms that acts transitively along the fibers of $T^*\mathcal{Q}$ *and* is compatible with the G-symmetries, in the sense that their ζ-action combines with G into an extended group. It begins by finding a linear representation of G on a vector space V such that at least one G-orbit $\mathcal{O}$ in V is diffeomorphic to $\mathcal{Q}$. Then, any dual vector $\alpha \in V^*$ can be naturally seen as an 1-form field on V, and it can be restricted to $\mathcal{O} \sim \mathcal{Q}$ to define a 1-form field on $\mathcal{Q}$. This 1-form field is exact because the function $f_\alpha : V \to \mathbb{R}$ defined by

$$f_\alpha(v) := \alpha(v) \tag{9.22}$$

satisfies $\alpha = df_\alpha$ (note the slight abuse of notation here, where the subscript α of f is an element of V^*, while α in the left-hand side is the corresponding 1-form field on V). The group of momentum translations generated by $\alpha \in V^*$ in this way acts transitively along the fibers of $T^*\mathcal{Q}$, i.e., given any p and p' on the same fiber, $\pi(p) = \pi(p')$, there is always $\alpha \in V^*$ such that $p' = p - \alpha = \zeta_\alpha(p)$. We shall refer to this realization of the configuration space as an orbit in a vector space as the *embedding realization.*

Combining the "horizontal" and "vertical" transformations discussed above, we get a transitive group of symmetries of the phase space with structure $\widetilde{G} = V^* \rtimes G$, acting on the phase space as

$$\Gamma_{(\alpha,g)}(p) := \delta^*_{g^{-1}} p - \alpha \tag{9.23}$$

The product rule on $\widetilde{G}$ is defined so that Γ composes appropriately,

$$(\alpha, g)(\alpha', g') = (\alpha + \delta^*_{g^{-1}} \alpha', gg') \tag{9.24}$$

where α and α' are seen as 1-form fields on $\mathcal{Q}$. More abstractly, $\alpha \mapsto \delta^*_{g^{-1}} \alpha$ corresponds to the dual representation of G on V^*, which can be denoted by $\alpha \mapsto g\alpha$, so

$$(\alpha, g)(\alpha', g') = (\alpha + g\alpha', gg') \tag{9.25}$$

where α and α' are seen as elements of V^*. This group, or an appropriate central extension of it, can be taken as the quantizing group.

Let us now compute the classical charges associated with the group $\widetilde{G} = V^* \rtimes G$. First we consider the G part of the group. If $\xi \in \mathfrak{g}$ is an element of the algebra of G, let $\widetilde{X}_\xi$ denote the vector field induced by ξ on $\widetilde{\mathcal{P}}$. The associated charge P_ξ, interpreted as a "momentum variable", is defined by

$$dP_\xi = -\imath_{\widetilde{X}_\xi}\omega = -\imath_{\widetilde{X}_\xi} d\theta = d[\theta(\widetilde{X}_\xi)] \tag{9.26}$$

where we have used that θ is invariant under the G action, so $\pounds_{\widetilde{X}_\xi}\theta = \imath_{\widetilde{X}_\xi} d\theta + d\imath_{\widetilde{X}_\xi}\theta = 0$. Up to an additive constant, we can therefore choose

$$P_\xi = p(X_\xi) \tag{9.27}$$

where $X_\xi := \pi_* \widetilde{X}_\xi$.[46] Second, we consider the V^* part of the group. Since V^* is a vector space, we can naturally identify it with is Lie algebra. If $\alpha \in V^*$, seen as an element of the Lie algebra of V^*, let Y_α denote the vector field induced by α on $\widetilde{\mathcal{P}}$. The associated charges Q_α, interpreted as a "configuration variable", is defined by

$$dQ_\alpha = -\imath_{Y_\alpha}\omega = -\imath_{Y_\alpha} d\theta = -\pounds_{Y_\alpha}\theta = -\left.\frac{d}{dt}\zeta^*_{t\alpha}\theta\right|_{t=0} = \pi^*\alpha = \pi^* df_\alpha = d(f_\alpha \circ \pi) \tag{9.28}$$

where we have used $\imath_{Y_\alpha}\theta = 0$, the result in (9.20), and the relation $\alpha = df_\alpha$. Thus, up to an additive constant, we can choose

$$Q_\alpha = f_\alpha \circ \pi \tag{9.29}$$

[46] Note that $\widetilde{X}_\xi$ does project nicely under π, as can be seen from the fact that we can alternatively define X_ξ as the vector field induced by ξ on $\mathcal{Q}$ (via the action of G on $\mathcal{Q}$). In fact, let $\widetilde{\phi}_p(g) = gp$ and $\phi_x(g) = gx$, where $p \in \widetilde{\mathcal{P}}$ and $x \in \mathcal{Q}$, then $\widetilde{X}_\xi := \widetilde{\phi}_{p*}\xi$ and let us define $X_\xi := \phi_{x*}\xi$; so because of the relation $\pi(gp) = g\pi(p)$, which translates into $\pi \circ \widetilde{\phi}_p = \phi_{\pi(p)}$, we have $X_\xi = \pi_* \widetilde{X}_\xi$.

Note that in the embedding realization $x \in \mathcal{Q}$ is seen a vector in V specifying a point in the orbit $\mathcal{O} \sim \mathcal{Q}$, so this expression for the charge, evaluated at $p \in \widetilde{\mathcal{P}}$, reads simply $Q_\alpha = \alpha(x)$, where $x = \pi(p)$.

Now let us discuss the general form of the algebra of those charges. Since $\widetilde{G}$ has a semi-direct product structure, its Lie algebra is a semi-direct sum $\widetilde{\mathfrak{g}} = V^* \oplus \mathfrak{g}$. If the elements of $\widetilde{\mathfrak{g}}$ are denoted by $(\alpha;\xi)$, the product rule derived from (9.24) is given by

$$[(\alpha;\xi),(\alpha';\xi')] = (\pounds_{X_{\xi'}}\alpha - \pounds_{X_\xi}\alpha';[\xi,\xi']) \tag{9.30}$$

in which $[\xi,\xi']$ is just the product rule in $\mathfrak{g}$. As we know, the Poisson algebra of the charges P_ξ and Q_α should form a representation up to central charges of the underlying symmetry algebra, so we can generically write

$$\begin{aligned}\{P_\xi, P_{\xi'}\} &= P_{[\xi,\xi']} + z(\xi,\xi')\\ \{Q_\alpha, P_\xi\} &= Q_{\pounds_{X_\xi}\alpha} + z(\alpha,\xi)\\ \{Q_\alpha, Q_{\alpha'}\} &= z(\alpha,\alpha')\end{aligned} \tag{9.31}$$

where $z(\xi,\xi')$, $z(\alpha,\xi)$ and $z(\alpha,\alpha')$ are the 2-cocycle elements (which are constants on the phase space).[47] To be explicit, let us write the charges computed above as

$$P_\xi = p(X_\xi) + c_\xi$$
$$Q_\alpha = f_\alpha \circ \pi + c_\alpha \tag{9.32}$$

where c_ξ and c_α are generic constants (on the phase space) depending linearly on ξ and α, respectively. We wish to derive a formula for the 2-cocycles z in terms of these additive constants c.

Let us start with the Q charges. Their Poisson brackets are

$$\{Q_\alpha, Q_{\alpha'}\} = -\omega(Y_\alpha, Y_{\alpha'}) = 0 \tag{9.33}$$

At the same time, $[\alpha, \alpha'] = 0$, so $Q_{[\alpha,\alpha']} = 0$, which gives

$$z(\alpha, \alpha') = 0 \tag{9.34}$$

regardless of c.

Now we consider the P charges. Their Poisson brackets are

$$\{P_\xi, P_{\xi'}\} = -\omega(\widetilde{X}_\xi, \widetilde{X}_{\xi'}) = -d\theta(\widetilde{X}_\xi, \widetilde{X}_{\xi'}) \tag{9.35}$$

[47] Note that we are using a simplified notation here, in which $z(\xi, \xi') := z((0;\xi),(0;\xi'))$, $z(\alpha, \xi) := z((\alpha;0),(0;\xi))$ and $z(\alpha, \alpha') := z((\alpha;0),(\alpha';0))$. In fact, this is what we mean whenever we write an element of $\mathfrak{g}$ or V^* in a place supposed to feature an element of $\widetilde{\mathfrak{g}}$.

which can be manipulated as

$$d\theta(\widetilde{X}_\xi, \widetilde{X}_{\xi'}) = -\imath_{\widetilde{X}_{\xi'}} d\theta(\widetilde{X}_\xi) = -\left(\pounds_{\widetilde{X}_{\xi'}}\theta - d\,\imath_{\widetilde{X}_{\xi'}}\theta\right)\widetilde{X}_\xi = d\left(\theta(\widetilde{X}_{\xi'})\right)\widetilde{X}_\xi = \widetilde{X}_\xi\left(\theta(\widetilde{X}_{\xi'})\right) \tag{9.36}$$

where the last term is read as the vector $\widetilde{X}_\xi$ deriving the scalar $\theta(\widetilde{X}_{\xi'})$. Thus

$$\widetilde{X}_\xi\left(\theta(\widetilde{X}_{\xi'})\right) = \pounds_{\widetilde{X}_\xi}\left(\theta(\widetilde{X}_{\xi'})\right) = \theta\left(\pounds_{\widetilde{X}_\xi}\widetilde{X}_{\xi'}\right) = \theta\left([\widetilde{X}_\xi, \widetilde{X}_{\xi'}]\right) \tag{9.37}$$

so we get

$$d\theta(\widetilde{X}_\xi, \widetilde{X}_{\xi'}) = \theta\left([\widetilde{X}_\xi, \widetilde{X}_{\xi'}]\right) \tag{9.38}$$

In addition,

$$\theta\left([\widetilde{X}_\xi, \widetilde{X}_{\xi'}]\right) = p\left(\pi_*[\widetilde{X}_\xi, \widetilde{X}_{\xi'}]\right) = p\left([X_\xi, X_{\xi'}]\right) = -p\left(X_{[\xi,\xi']}\right) \tag{9.39}$$

Hence,

$$\{P_\xi, P_{\xi'}\} = p\left(X_{[\xi,\xi']}\right) \tag{9.40}$$

which leads to the 2-cocycle

$$z(\xi, \xi') = -c_{[\xi,\xi']} \tag{9.41}$$

We see that we can eliminate this 2-cocycle by choosing $c_\xi = 0$.

Next we consider the mixed Q and P brackets, which give

$$\{Q_\alpha, P_\xi\} = -\omega(Y_\alpha, \widetilde{X}_\xi) = -d\theta(Y_\alpha, \widetilde{X}_\xi) = -\imath_{Y_\alpha} d\theta(\widetilde{X}_\xi) = \pi^* \alpha(\widetilde{X}_\xi) = \alpha(X_\xi) \quad (9.42)$$

and, since this should be a function on the phase space, it should be more precisely written as $\alpha(X_\xi) \circ \pi$. Thus,

$$z(\alpha, \xi) = \alpha(X_\xi) \circ \pi - f_{\pounds_{X_\xi} \alpha} \circ \pi - c_{\pounds_{X_\xi} \alpha} \quad (9.43)$$

This cocycle may be non-trivial, i.e., it may not be removable by a choice of c_α. In that case, we must centrally extend $\widetilde{G}$ by this 2-cocycle, so that the Lie algebra of the new group is isomorphic to the Poisson algebra of the new charges.

10 The canonical group for the diamond

Now we apply Isham's scheme to the reduced phase space of the diamond, $\widetilde{\mathcal{P}} = T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$. Being a cotangent bundle over a homogeneous space, $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, we may try to construct the quantizing group for this phase space according to the algorithm in the previous section. The natural choice for the group G here is $\mathit{Diff}^+(S^1)$, acting on the left of $\mathcal{Q}$ as

$$\psi'[\psi] := [\psi' \circ \psi] \quad (10.1)$$

where $[\psi] = q(\psi)$ is the quotient by the right $PSL(2,\mathbb{R})$ action. The definition above makes sense because the quotient acts from the right, so if $\chi \in PSL(2,\mathbb{R})$, $\psi'[\psi\chi] = [\psi'\psi\chi] = [\psi'\psi] = \psi'[\psi]$. The next step is to find a representation of $\mathit{Diff}^+(S^1)$, on a vector space V, such that at least one G-orbit $\mathcal{O}$ in V is isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. However, we are not aware of a representation of $\mathit{Diff}^+(S^1)$ which has this property. On the other hand, it is well known that the coadjoint representation of the Virasoro group does have an orbit isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ [94, 95, 96, 97]. Moreover, it is just as natural to choose G as the Virasoro group, *Vira*, given that it is a central extension of $\mathit{Diff}^+(S^1)$ and we can simply define its action on $\mathcal{Q}$ such that the central element acts trivially.

10.1 The Virasoro group

Let us begin by recalling the definition of the Virasoro group and its Lie algebra. (For reference see, e.g., [98, 99, 100, 101, 45].) The Virasoro group, *Vira*, is a one-dimensional central extension of $\mathit{Diff}^+(S^1)$ which can be defined from its Lie algebra, the Virasoro algebra, $\mathfrak{vira} = \widehat{\mathfrak{diff}(S^1)}$, a central extension of $\mathfrak{diff}(S^1)$. An element $\widehat{\xi}$ of the Virasoro algebra is characterized by an element ξ of $\mathfrak{diff}(S^1)$ (which is identified with a vector field $\xi(\theta)\partial_\theta$ on S^1), together with a central element component, $x\widehat{c}$, where $x \in \mathbb{R}$ and $\widehat{c}$ is a vector in the center of $\mathfrak{vira}$. Thus, we write

$$\widehat{\xi} = \xi(\theta)\partial_\theta + x\widehat{c} \in \mathfrak{vira} \tag{10.2}$$

Sometimes we may also write $\widehat{\xi} = \xi + x\widehat{c}$; we hope that it will be clear from the context whether ξ is referring to the vector field $\xi(\theta)\partial_\theta$ or just the real function $\xi(\theta)$. The product rule in $\mathfrak{vira}$ is defined by

$$[\xi\partial_\theta + x\widehat{c}, \eta\partial_\theta + y\widehat{c}] = \left(\eta\xi' - \xi\eta'\right)\partial_\theta + \widehat{c}\int d\theta\left(\eta\xi''' - \xi\eta'''\right) \tag{10.3}$$

where the prime denotes derivative with respect to θ. Note that the ∂_θ component in the right-hand side is *equal* to the product of $\xi\partial_\theta$ and $\eta\partial_\theta$ in the algebra $\mathfrak{diff}(S^1)$, which is *minus* the vector field brackets of these fields on S^1 (see the "$\mathfrak{diff}(S^1)$" entry in App. A). Also, note that $\widehat{c}$ spans the center of the algebra.

The global topology of the Virasoro group, defined as the exponentiation of $\mathfrak{vira}$, is not completely fixed. We shall take it to have topology $\mathit{Diff}^+(S^1) \times \mathbb{R}$, so that it can be characterized by pairs

$$\widehat{\psi} = (\psi, r) \in \mathit{Vira} \tag{10.4}$$

where $\psi \in \mathit{Diff}^+(S^1)$ and $r \in \mathbb{R}$. Note that an alternative option would be to define the exponentiated group as being simply connected. This corresponds to "unwrapping" the $\mathit{Diff}^+(S^1)$ factor (i.e., taking its universal cover), given that $\mathit{Diff}^+(S^1)$ has fundamental group $\mathbb{Z}$ due to its $SO(2)$ subgroup of rotations (see App. D). Later, when considering quantizations based on projective representations, we will discuss its universal cover (denoted by $\underline{\mathit{Vira}}$).

The Lie algebra $\mathfrak{vira}$ can be exponentiated into the group *Vira* with the help of Baker-Campbell-Hausdorff formula, which reads

$$\exp(\widehat{\xi})\exp(\widehat{\eta}) = \exp\left(\widehat{\xi} + \widehat{\eta} + \frac{1}{2}[\widehat{\xi}, \widehat{\eta}] + \frac{1}{12}[\widehat{\xi}, [\widehat{\xi}, \widehat{\eta}]] - \frac{1}{12}[\widehat{\eta}, [\widehat{\xi}, \widehat{\eta}]] + \cdots\right) \tag{10.5}$$

where $\widehat{\xi}$, $\widehat{\eta} \in \mathfrak{vira}$ and the "$\cdots$" consists of higher-order algebra products. This formula allows us to reconstruct the group, at least in a neighborhood of the identity.[48] From the Baker-Campbell-Hausdorff formula, it is clear that when restricting to $\widehat{\xi} = \xi\partial_\theta$ the exponential map yields $\mathit{Diff}^+(S^1)$, and when restricting to $\widehat{\xi} = x\widehat{c}$ the exponential map yields $(\mathbb{R}, +)$. So let us define

$$(\psi_\xi, x) := \exp(\xi\partial_\theta + x\widehat{c}) \tag{10.6}$$

in which

$$\psi_\xi := \exp_{\mathfrak{diff}}(\xi\partial_\theta) \tag{10.7}$$

where $\exp_{\mathfrak{diff}}$ denotes the exponential from $\mathfrak{diff}(S^1)$ into $\mathit{Diff}^+(S^1)$. Using that $\widehat{c}$ is a central element, the Baker-Campbell-Hausdorff formula also gives

$$\exp(\xi\partial_\theta)\exp(x\widehat{c}) = \exp(\xi\partial_\theta + x\widehat{c}) \tag{10.8}$$

[48] Rigorously speaking, this is only true for finite-dimensional algebras, since in this case it can be shown that the exponential map yields a diffeomorphism between a neighborhood of 0 in the algebra and a neighborhood of e in the group. In fact, the exponential map from $\mathfrak{diff}(S^1)$ into $\mathit{Diff}^+(S^1)$ is not surjective, even in a neighborhood of the identity. This happens because there are diffeomorphims ψ, arbitrarily close to the identity, that do not admit a "square root", i.e., there is no ϕ such that $\psi = \phi^2$. If ψ were equal to the exponential of some algebra element, $\psi = \exp(\xi)$, then $\phi = \exp(\xi/2)$ would be the square root of ψ.

which in the notation above reads

$$(\psi_\xi, x) = (\psi_\xi, 0)(I, x) \tag{10.9}$$

With a natural extrapolation we can pose

$$(\psi, r) = (\psi, 0)(I, r) \tag{10.10}$$

for all $\psi \in \mathit{Diff}^+(S^1)$ and $r \in \mathbb{R}$, which implies that every element of *Vira* can be decomposed as a product of an element of $\mathit{Diff}^+(S^1) \subset \mathit{Vira}$ and an element of the center $\mathbb{R} \subset \mathit{Vira}$.

Next we derive the general form for the product rule in *Vira*. First, write (10.3) in the compact form $[\xi\partial_\theta + x\widehat{c}, \eta\partial_\theta + y\widehat{c}] = [\xi\partial_\theta, \eta\partial_\theta]_{\mathfrak{diff}} + w(\xi,\eta)\widehat{c}$, where $w(\xi,\eta)$ is the bilinear functional of ξ and η given by

$$w(\xi, \eta) := \int d\theta \left(\eta\xi''' - \xi\eta'''\right) \tag{10.11}$$

From Baker-Campbell-Hausdorff formula we get,

$$\begin{aligned}
&\exp(\xi\partial_\theta + x\widehat{c})\exp(\eta\partial_\theta + y\widehat{c}) = \\
&= \exp\left[\left(\xi + \eta + \frac{1}{2}[\xi,\eta]_{\mathfrak{diff}} + \cdots\right)\partial_\theta + \left(x + y + \frac{1}{2}w(\xi,\eta) + \frac{1}{12}w(\xi,[\xi,\eta]_{\mathfrak{diff}}) + \cdots\right)\widehat{c}\right] \\
&= \exp\left[\left(\xi + \eta + \frac{1}{2}[\xi,\eta]_{\mathfrak{diff}} + \cdots\right)\partial_\theta\right]\exp\left[(x + y + W(\xi,\eta))\,\widehat{c}\right] \qquad (10.12)
\end{aligned}$$

where $W(\xi,\eta) := \frac{1}{2}w(\xi,\eta) + \frac{1}{12}w(\xi,[\xi,\eta]_{\mathfrak{diff}}) + \cdots$ is a (real) function of $\xi, \eta \in \mathfrak{diff}(S^1)$. The first factor can be identified as the product rule in *Diff*$^+(S^1)$, while the second factor gives a central element,

$$\exp(\xi\partial_\theta + x\widehat{c})\exp(\eta\partial_\theta + y\widehat{c}) = (\psi_\xi\psi_\eta, 0)(I, x+y+W(\xi,\eta)) = (\psi_\xi\psi_\eta, x+y+W(\xi,\eta)) \tag{10.13}$$

Since ξ determines ψ_ξ, we can by a slight abuse of notation write $W(\xi,\eta)$ as $W(\psi_\xi,\psi_\eta)$. The general product rule in *Vira* then reads

$$(\psi,a)(\phi,b) = (\psi\phi, a+b+W(\psi,\phi)) \tag{10.14}$$

Note that this derivation only defines W when acting on diffeomorphisms in the image of the exponential map (see footnote 48), but we can smoothly extend the domain of W so as to include the exceptional points. This function is called the *Bott 2-cocycle*, and it can be explicitly expressed as $W(\psi,\phi) = \int_{S^1}\log(\psi\circ\phi)' d\log\phi'$, where the prime denotes derivative with respect to the angle [101].

10.1.1 Adjoint representation and orbits

Let us now consider the adjoint representation of the Virasoro group on its Lie algebra, and the corresponding adjoint representation of the Lie algebra on itself. As always, the adjoint representation of the Lie algebra on itself is given by the algebra product,

$$\mathrm{ad}_{\xi\partial_\theta + x\widehat{c}}(\eta\partial_\theta + y\widehat{c}) = \left(\eta\xi' - \xi\eta'\right)\partial_\theta + \widehat{c}\int d\theta\,\left(\eta\xi''' - \xi\eta'''\right) \tag{10.15}$$

As the adjoint action of the algebra on itself also corresponds to the "derivative" of adjoint action of the group on the algebra, we can think of $\mathrm{ad}_{\widehat{\xi}}\widehat{\eta}$ as the infinitesimal change of $\widehat{\eta}$ as one deform the identity in the group along a curve tangent to $\widehat{\xi}$ (see the "Ad_g, ad_g and ad_ξ" entries in App. A). Therefore we can "integrate" the action above to obtain the adjoint representation of the group. First note that $\widehat{c}$ is represented trivially, i.e., $\mathrm{ad}_{\widehat{c}} = 0$. Upon integration, it is clear that (I, r) will also be represented trivially, i.e., $\mathrm{ad}_{(I,r)} = 1$, where 1 denotes the identity operator on $\mathfrak{vira}$. Using the decomposition property in (10.10), and the fact that the adjoint action forms a representation of the group, we have

$$\mathrm{ad}_{(\psi,r)} = \mathrm{ad}_{(\psi,0)(I,r)} = \mathrm{ad}_{(\psi,0)}\mathrm{ad}_{(I,r)} = \mathrm{ad}_{(\psi,0)} \tag{10.16}$$

that is, the adjoint action does not depend on the central component of $\widehat{\psi} = (\psi, r)$. Accordingly, we can simplify the notation and just write ad_ψ instead of $\mathrm{ad}_{\widehat{\psi}}$, while keeping in mind that this refers to the adjoint representation of *Vira*, not *Diff*$^+(S^1)$. Note also that, since $\mathrm{ad}_{\widehat{\xi}}\widehat{c} = 0$, the adjoint map should act trivially on $\widehat{c}$, i.e., $\mathrm{ad}_\psi\widehat{c} = \widehat{c}$. Thus we have,

$$\mathrm{ad}_\psi\widehat{\eta} = \mathrm{ad}_\psi(\eta\partial_\theta) + y\widehat{c} \tag{10.17}$$

and we can focus on $\mathrm{ad}_\psi(\eta\partial_\theta)$. Note that $\mathrm{ad}_\xi(\eta\partial_\theta)$ changes both the ∂_θ and the central components of η. Since the way it changes the ∂_θ component is exactly the same as the action of the adjoint map in *Diff*$^+(S^1)$, we conclude that the "integrated" action must have the form

$$\mathrm{ad}_\psi(\eta\partial_\theta) = \psi_*(\eta\partial_\theta) + \Lambda_\psi(\eta)\widehat{c} \tag{10.18}$$

where Λ is some ψ-dependent linear functional of η. We can get a clue about what is Λ_ψ by computing the product of two adjoint maps,

$$\mathrm{ad}_\psi\mathrm{ad}_\phi\eta = \mathrm{ad}_\psi(\phi_*\eta + \Lambda_\phi(\eta)\widehat{c}) = \psi_*\phi_*\eta + \left(\Lambda_\psi(\phi_*\eta) + \Lambda_\phi(\eta)\right)\widehat{c} \tag{10.19}$$

where, for notational simplicity, η is being written in place of $\eta(\theta)\partial_\theta$. But since the adjoint map forms a representation of the group,

$$\mathrm{ad}_\psi\mathrm{ad}_\phi = \mathrm{ad}_{(\psi,0)}\mathrm{ad}_{(\phi,0)} = \mathrm{ad}_{(\psi,0)(\phi,0)} = \mathrm{ad}_{(\psi\phi,W(\psi,\phi))} = \mathrm{ad}_{\psi\phi} \tag{10.20}$$

which gives

$$\mathrm{ad}_\psi\mathrm{ad}_\phi\eta = (\psi\phi)_*\eta + \Lambda_{\psi\phi}(\eta)\widehat{c} \tag{10.21}$$

Hence we conclude that Λ must satisfy the relation

$$\Lambda_{\psi\phi} = \Lambda_\psi \circ \phi_* + \Lambda_\phi \tag{10.22}$$

A solution to this relation can be obtained from the *Schwarzian derivative* [100], a map from $Diff^+(S^1)$ into $C^\infty(S^1, \mathbb{R})$ defined as

$$S[\psi](\theta) := \frac{\psi'''(\theta)}{\psi'(\theta)} - \frac{3}{2}\left(\frac{\psi''(\theta)}{\psi'(\theta)}\right)^2 \tag{10.23}$$

in which ψ is being realized as a monotonous map from $[0, 2\pi]$ into $\mathbb{R}$ satisfying $\psi(2\pi) = \psi(0) + 2\pi$. It satisfies the relation

$$S[\psi \circ \phi] = (S[\psi] \circ \phi)\,(\phi')^2 + S[\phi] \tag{10.24}$$

that is, $S[\psi\phi](\theta) = S[\psi](\phi(\theta))\phi'(\theta)^2 + S[\phi](\theta)$. Comparing with (10.22), we infer

$$\Lambda_\psi(\eta) = \kappa \int d\theta\, S[\psi](\theta)\, \eta(\theta) \tag{10.25}$$

where κ is a coefficient to be adjusted. If $\bar{\theta} = \phi(\theta)$, we have

$$(\bar{\eta}\partial_\theta)\big|_{\bar{\theta}} := \phi_*(\eta\partial_\theta) = \eta(\theta)\phi'(\theta)\partial_\theta \tag{10.26}$$

that is, $\bar{\eta}(\bar{\theta}) = \phi'(\theta)\eta(\theta)$. Then,

$$\begin{aligned}
\Lambda_{\psi\phi}(\eta) &= \kappa \int d\theta\, S[\psi\phi](\theta)\, \eta(\theta) \\
&= \kappa \int d\theta\, \left(S[\psi](\phi(\theta))\phi'(\theta)^2 + S[\phi](\theta) \right) \eta(\theta) \\
&= \kappa \int d\theta\, \phi'(\theta) S[\psi](\phi(\theta))\phi'(\theta)\eta(\theta) + \Lambda_\phi(\eta) \\
&= \kappa \int d\bar{\theta}\, S[\psi](\bar{\theta}))\bar{\eta}(\bar{\theta}) + \Lambda_\phi(\eta) \\
&= \Lambda_\psi(\phi_* \eta) + \Lambda_\phi(\eta)
\end{aligned} \tag{10.27}$$

where in the second line we used property (10.24) and in the fourth line we changed the integration variable from θ to $\bar{\theta}$. Finally, in order to determine κ, we compute the derivative of (10.17) and demand that it matches with (10.15). If $t \mapsto \psi_t$ is a curve in *Diff*$^+(S^1)$ tangent to $\xi \in \mathfrak{diff}(S^1)$, then we have

$$\mathrm{ad}_\xi(\eta) = \frac{d}{dt}\mathrm{ad}_{\psi_t}(\eta) = [\xi, \eta]_{\mathfrak{diff}} + \widehat{c}\kappa \int d\theta\, \frac{d}{dt} S[\psi_t](\theta)\, \eta(\theta) \tag{10.28}$$

where the derivative is evaluated at $t = 0$. If $\xi = \xi(\theta)\partial_\theta$, then a possible choice for the curve is $\psi_t(\theta) = \theta + t\xi(\theta)$, so that $S[\psi_t](\theta) = t\xi'''(\theta) + \mathcal{O}(t^2)$. Thus,

$$\mathrm{ad}_\xi(\eta) = \left(\eta\xi' - \xi\eta' \right) \partial_\theta + \widehat{c}\kappa \int d\theta\, \xi''' \eta \tag{10.29}$$

Note that, by integrating by parts, we have $\int d\theta\, \xi'''\eta = -\int d\theta\, \xi\eta'''$, and so we see that the choice $\kappa = 2$ reproduces (10.15). That is,

$$\mathrm{ad}_{\widehat{\psi}}\widehat{\eta} = \psi_*(\eta\partial_\theta) + \widehat{c}\left(y + 2\int d\theta\, S[\psi](\theta)\,\eta(\theta)\right) \tag{10.30}$$

where $\widehat{\psi} = (\psi, r)$ and $\widehat{\eta} = \eta\partial_\theta + y\widehat{c}$.

Before moving on to the coadjoint representation, let us show that there are no orbits of the adjoint action which are isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. To do so, define the *little algebra* of $\widehat{\eta} \in \mathfrak{vira}$, with respect to the adjoint action, as

$$\mathfrak{h}_{\mathrm{ad}}(\widehat{\eta}) := \{\widehat{\xi} \in \mathfrak{vira}\,,\ \mathrm{ad}_{\widehat{\xi}}\widehat{\eta} = 0\} \tag{10.31}$$

The *little group* of the orbit of $\widehat{\eta}$ under the adjoint action is the exponentiation of the little algebra $\mathfrak{h}_{\mathrm{ad}}(\widehat{\eta})$. Since $\widehat{c}$ is adjointly represented as the zero operator, it always belongs to the little algebra. If there exits some orbit isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ then there must exist some little group isomorphic to $PSL(2,\mathbb{R}) \times \mathbb{R}$, where $\mathbb{R}$ refers to the center of *Vira*, and so there must exist some little algebra isomorphic to $\mathfrak{psl}(2,\mathbb{R}) \oplus \widehat{c}$. As we shall see, however, there is no such little algebra for the adjoint action. Note that a necessary condition for $\widehat{\xi} \in \mathfrak{h}_{\mathrm{ad}}(\widehat{\eta})$ is

$$\eta\xi' - \xi\eta' = 0 \tag{10.32}$$

so in every open interval of S^1 in which $\eta \neq 0$ we have $\xi = k\eta$ for some constant k.

This condition is therefore also sufficient, since it implies $\eta\xi''' - \xi\eta''' = 0$. There are three cases to consider,

(i) If η vanishes only at isolated points, and at least one of its derivatives do not vanish at each of those points, then by requiring that ξ is smooth we get that $\xi = k\eta$ on the whole S^1. In this case the little algebra is two-dimensional, $\mathfrak{h}_{\mathrm{ad}}(\widehat{\eta}) = \eta \oplus \widehat{c}$;

(ii) If η vanishes only at isolated points, and all of its derivatives also vanish at n of those points (called "flat points"), then the smoothness of ξ is not sufficient to relate k's at different sections R_i of S^1, where each R_i is the closed interval between two consecutive flat points. In this case, ξ can be piecewise-defined as $\xi(\theta) = k_i\eta(\theta)$, for $\theta \in R_i$, and so the little algebra is $(n+1)$-dimensional, $\mathfrak{h}_{\mathrm{ad}}(\widehat{\eta}) = (\eta \oplus \eta \cdots \oplus \eta) \oplus \widehat{c}$;

(iii) If η vanishes in an open set of S^1, then ξ is not restricted on that interval and the little algebra is infinite-dimensional;

Note that cases (i) and (iii) do not have the right dimension to be isomorphic to $\mathfrak{psl}(2,\mathbb{R}) \oplus \widetilde{c}$, but case (ii) can have 4 dimensions if $n = 3$. However, it does not have the right algebraic structure, since it is commutative.

10.1.2 Coadjoint representation and orbits

Now we consider the coadjoint representation of *Vira* on its dual Lie algebra, $\mathfrak{vira}^*$, and the corresponding coadjoint representation of $\mathfrak{vira}$ on $\mathfrak{vira}^*$. Similar to our manner of characterizing elements of $\mathfrak{diff}^*(S^1)$, we characterize elements of $\mathfrak{vira}^*$ as $\widetilde{\alpha} = \alpha + \alpha_0\widetilde{c}$ where $\alpha := \alpha(\theta)d\theta^2$ is a quadratic form on S^1 and $\widetilde{c}$ is dual to $\widehat{c}$ in the sense that $\widetilde{c}(\eta + y\widehat{c}) = y$. The pairing of $\widetilde{\alpha} \in \mathfrak{vira}^*$ with $\widehat{\eta} \in \mathfrak{vira}$ is defined as

$$\widetilde{\alpha}(\widehat{\eta}) := \int_{S^1} \alpha(\eta) + \alpha_0 y = \int d\theta\, \alpha(\theta)\eta(\theta) + \alpha_0 y \tag{10.33}$$

From the definition of the coadjoint action (see "coad_g" entry in App. A) we have

$$\mathrm{coad}_{(\psi,r)}\widetilde{\alpha}\,(\widehat{\eta}) = \widetilde{\alpha}\left(\mathrm{ad}_{(\psi,r)^{-1}}\widehat{\eta}\right) = \widetilde{\alpha}\left(\mathrm{ad}_{(\psi^{-1},-r-W(\psi,\psi^{-1}))}\widehat{\eta}\right) \tag{10.34}$$

which from (10.30) gives

$$\begin{aligned} \mathrm{coad}_{(\psi,r)}\widetilde{\alpha}\,(\widehat{\eta}) &= \widetilde{\alpha}\left(\psi_*^{-1}\eta + \widehat{c}\left(y + 2\int d\theta\, S[\psi^{-1}](\theta)\,\eta(\theta)\right)\right) \\ &= \int \alpha(\psi_*^{-1}\eta) + 2\alpha_0 \int d\theta\, S[\psi^{-1}](\theta)\,\eta(\theta) + \alpha_0 y \\ &= \left(\psi_*(\alpha d\theta^2) + 2\alpha_0 S[\psi^{-1}]d\theta^2 + \alpha_0\widetilde{c}\right)(\widehat{\eta}) \end{aligned} \tag{10.35}$$

where, as usual, $\psi_* = \psi^{-1*}$. Therefore,

$$\mathrm{coad}_{(\psi,r)}\widetilde{\alpha} = \psi_*(\alpha d\theta^2) + 2\alpha_0 S[\psi^{-1}]d\theta^2 + \alpha_0\widetilde{c} \tag{10.36}$$

Note that r also acts trivially through the coadjoint map, so we can use the shortened notation coad_ψ. For later reference, we can evaluate the push-forward explicitly to get

$$\operatorname{coad}_\psi(\alpha d\theta^2 + \alpha_0 \widetilde{c})\Big|_\theta = \frac{\alpha(\psi^{-1}(\theta)) - 2\alpha_0 S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} d\theta^2 + \alpha_0 \widetilde{c} \tag{10.37}$$

where we have used (10.24) to relate $S[\psi^{-1}]$ with $S[\psi]$. The coadjoint representation of the group leads to a coadjoint representation of the algebra, obtained by differentiation. The simplest way to derive this representation is by using directly the adjoint representation of the algebra, given in (10.15), since the coadjoint action is (minus) the dual of the adjoint action. We have,

$$\begin{aligned}
\operatorname{coad}_{\widehat{\xi}}\widetilde{\alpha}(\widehat{\eta}) = -\widetilde{\alpha}(\operatorname{ad}_{\widehat{\xi}}\widehat{\eta}) &= -\widetilde{\alpha}\left(\left(\eta\xi' - \xi\eta'\right)\partial_\theta + \widehat{c}\int d\theta\,\left(\eta\xi''' - \xi\eta'''\right)\right) \\
&= -\int d\theta\,\alpha\left(\eta\xi' - \xi\eta'\right) - \alpha_0\int d\theta\,\left(\eta\xi''' - \xi\eta'''\right) \\
&= -\int d\theta\,\eta\left(2\alpha\xi' + \xi\alpha' + 2\alpha_0\xi'''\right)
\end{aligned} \tag{10.38}$$

from which we read

$$\operatorname{coad}_{\widehat{\xi}}\widetilde{\alpha} = -\left(2\alpha\xi' + \xi\alpha' + 2\alpha_0\xi'''\right)d\theta^2 \tag{10.39}$$

Naturally, this can also be obtained by explicitly differentiating (10.37).

To show that there is a coadjoint orbit of *Vira* which is isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, it is convenient to define the little algebra of $\widetilde{\alpha} \in \mathfrak{vira}^*$, with respect to the coadjoint action, as

$$\mathfrak{h}_{\text{coad}}(\widetilde{\alpha}) := \{\widehat{\xi} \in \mathfrak{vira}\,,\ \text{coad}_{\widehat{\xi}}\widetilde{\alpha} = 0\} \tag{10.40}$$

Note that the central element, $\widehat{c}$, is always in h_{coad}. Let us particularize to the little algebra of $\widetilde{\alpha}_0 := \alpha_0 d\theta^2 + a\widetilde{c}$, where $\alpha_0 \in \mathbb{R}$ is a constant. The equation for $\widehat{\xi} \in \mathfrak{h}_{\text{coad}}(\alpha_0 d\theta^2 + a\widetilde{c})$ reads

$$\alpha_0 \xi' + a\xi''' = 0 \tag{10.41}$$

This differential equation (of the third order) has a general solution of the form

$$\xi = h_0 + h_1 \cos w\theta + h_2 \sin w\theta \tag{10.42}$$

where h_0, h_1 and h_2 are integration constants and $w = \sqrt{\alpha_0/a}$. Note that if $w \in \mathbb{Z}$ then all terms are allowed, but otherwise only the h_0 term is allowed (since ξ must be 2π-periodic so as to be smooth on S^1). In particular, if $\alpha_0 = a$, so that $w = 1$, then the little algebra is spanned by ∂_θ, $\cos\theta\,\partial_\theta$, $\sin\theta\,\partial_\theta$ and $\widehat{c}$, which is precisely $\mathfrak{psl}(2,\mathbb{R}) \oplus \widehat{c}$. Hence, the coadjoint orbit of

$$\widetilde{\varepsilon} = d\theta^2 + \widetilde{c} \tag{10.43}$$

is isomorphic to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. Rigorously speaking, this statement is speculative since we have only considered the little algebra, and thus we would like to confirm that the *little group* of $\widetilde{\varepsilon}$ is indeed $PSL(2,\mathbb{R}) \times \mathbb{R}$. The little group of

$\widetilde{\alpha} \in \mathfrak{vira}^*$, with respect to the coadjoint action, is defined as

$$H_{\text{coad}}(\widetilde{\alpha}) := \{(\psi, r) \in \mathit{Vira}\,,\ \text{coad}_{(\psi,r)}\widetilde{\alpha} = \widetilde{\alpha}\} \tag{10.44}$$

Naturally the center of *Vira*, (I, r), is always in the little group. Given formula (10.37) for the coadjoint action, particularized to $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$, we have

$$\text{coad}_\psi \widetilde{\varepsilon}\Big|_\theta = \frac{1 - 2S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} d\theta^2 + \widetilde{c} \tag{10.45}$$

so the condition that ψ is in $H_{\text{coad}}(\widetilde{\varepsilon})$ translates into

$$S[\psi](\theta) = \frac{1 - \psi'(\theta)^2}{2} \tag{10.46}$$

We now wish to show that $PSL(2,\mathbb{R}) \subset H_{\text{coad}}(\widetilde{\varepsilon})$. Recall that $PSL(2,\mathbb{R}) \subset \mathit{Diff}^+(S^1)$ is defined as the boundary action of the group of holomorphic automorphisms of the complex unit disc $\mathbb{D} \subset \mathbb{C}$. That is, if $\psi \in PSL(2,\mathbb{R}) \subset \mathit{Diff}^+(S^1)$, then there exists a holomorphic function $f : \mathbb{D} \to \mathbb{D}$ such that $f(e^{i\theta}) = e^{i\psi(\theta)}$. In fact, f is a Mobius transformation with the general form

$$f(z) = e^{ib} \frac{z - a}{1 - \bar{a}z} \tag{10.47}$$

where $b \in \mathbb{R}$ and $a \in \mathbb{D}$ (see App. B). From the chain rule we have

$$\psi' = \frac{z}{f}\frac{df}{dz} \tag{10.48}$$

in which $z = e^{i\theta}$ is at the boundary of the disc, $\partial\mathbb{D} \sim S^1$. From this, we can show that

$$S[\psi] := \frac{\psi'''}{\psi'} - \frac{3}{2}\left(\frac{\psi''}{\psi'}\right)^2 = \frac{1-\psi'^2}{2} - z^2\mathfrak{S}[f] \tag{10.49}$$

where $\mathfrak{S}[f]$ is the Schwarzian derivative for holomorphic functions, defined analogously to (10.22) by

$$\mathfrak{S}[f](z) := \frac{d^3f/dz^3}{df/dz} - \frac{3}{2}\left(\frac{d^2f/dz^2}{df/dz}\right)^2 \tag{10.50}$$

This derivative has the property that $\mathfrak{S}[f] = 0$ if and only if f is a Mobius transformation. Therefore $\mathfrak{S}[f]$ vanishes for the $PSL(2,\mathbb{R})$ transformations, and relation (10.49) implies that ψ satisfies (10.46), establishing that $PSL(2,\mathbb{R})$ is indeed in the little group of $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$.[49]

[49] Note that only inclusion has been proven, $PSL(2,\mathbb{R}) \times \mathbb{R} \subset H_{\text{coad}}(\widetilde{\varepsilon})$, not equality. The fact that the little algebra is $\mathfrak{psl}(2,\mathbb{R}) \oplus \widehat{c}$ implies that $H_{\text{coad}}(\widetilde{\varepsilon})$ cannot be a higher-dimensional extension of $PSL(2,\mathbb{R}) \times \mathbb{R}$. While in principle it could be a discrete extension, $PSL(2,\mathbb{R})$ cannot be discretely extended within $\mathit{Diff}^+(S^1)$.

10.2 The canonical group for the reduced phase space

In view of the result above, and Isham's algorithm discussed in Sec. 9.3, it is possible to construct the quantum theory based on the group

$$\widetilde{G} = (\mathfrak{vira}^*)^* \rtimes \mathit{Vira} \tag{10.51}$$

in which the Virasoro group acts as the generator of "configuration translations" on $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ and the abelian group $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$ (whose product rule is given by vector addition, $\widehat{\eta}\widehat{\eta}' := \widehat{\eta} + \widehat{\eta}'$) acts as the generator of "momentum translations" along the fibers of $T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$.

We now describe explicitly how this group acts on the phase space. A point $[\psi]$ in the configuration space can be conveniently characterized by giving a Virasoro element that maps $d\theta^2 + \widetilde{c}$ into $[\psi]$. Since the central elements (I, r) act trivially, we need only to use elements of $\mathit{Diff}^+(S^1)$ in this characterization. In other words, the coadjoint representation of the Virasoro group provides a explicit projection map $\psi \mapsto [\psi]$, from $\mathit{Diff}^+(S^1)$ to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, given by

$$[\psi] := \mathrm{coad}_\psi(d\theta^2 + \widetilde{c}) = \frac{1 - 2S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} d\theta^2 + \widetilde{c} \tag{10.52}$$

The coadjoint action on $[\psi]$ is, as expected, just the right-action by $\mathit{Diff}^+(S^1)$ proposed in (10.1),

$$\phi[\psi] := \mathrm{coad}_\phi[\psi] = \mathrm{coad}_\phi \mathrm{coad}_\psi(d\theta^2 + \widetilde{c}) = \mathrm{coad}_{\phi\psi}(d\theta^2 + \widetilde{c}) = [\phi\psi] \tag{10.53}$$

so that we have $\delta_{(\phi,r)} = \mathrm{coad}_\phi$. Elements of the group will be denoted by

$$(\widehat{\eta}; \widehat{\phi}) \in \widetilde{G} \tag{10.54}$$

where $\widehat{\eta} \in \mathfrak{vira}$ given the natural identification $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$. The action of $\widetilde{G}$ on the phase space is, according to (9.23),

$$\Gamma_{(\widehat{\eta};\widehat{\phi})}(p) = \delta^*_{\widehat{\phi}^{-1}} p - \widehat{\eta} = \mathrm{coad}^*_{\widehat{\phi}^{-1}} p - \widehat{\eta} \tag{10.55}$$

in which, p, a 1-form at $[\psi] \in \mathcal{Q} \subset \mathfrak{vira}^*$, can be seen as the restriction of an element of $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$, naturally identified with a 1-form field on $\mathfrak{vira}^*$, to $\mathcal{Q}$. In this view, $\mathrm{coad}^*_{\widehat{\phi}^{-1}}$ is simply the dual representation of coad on $(\mathfrak{vira}^*)^*$, which is naturally identified with the adjoint representation on $\mathfrak{vira}$, so that

$$\Gamma_{(\widehat{\eta},\widehat{\phi})}(p) = \mathrm{ad}_\phi p - \widehat{\eta} \tag{10.56}$$

The group product rule then reads,

$$(\widehat{\xi}; \widehat{\psi})(\widehat{\xi}', \widehat{\psi}') = (\widehat{\xi} + \mathrm{ad}_\psi \widehat{\xi}', \widehat{\psi}\widehat{\psi}') \tag{10.57}$$

so, evidently, the semi-direct product is defined with respect to the adjoint action of *Vira* on $\mathfrak{vira}$.

The algebra of the group, $\widetilde{\mathfrak{g}} = \mathfrak{vira}^c \oplus \mathfrak{vira}$, is given, as in (9.30), by

$$[(\widehat{\eta};\widehat{\xi}),(\widehat{\eta}';\widehat{\xi}')] = (\pounds_{X_{\widehat{\xi}'}}\widehat{\eta} - \pounds_{X_{\widehat{\xi}}}\widehat{\eta}';[\widehat{\xi},\widehat{\xi}']) \tag{10.58}$$

where $[\widehat{\xi},\widehat{\xi}']$ is the product in $\mathfrak{vira}$. Note that the abelian factor, associated with the momentum translations, is denoted as $\mathfrak{vira}^c$ to emphasize that, while this factor is isomorphic to $\mathfrak{vira}$ as a vector space, it is not the usual $\mathfrak{vira}$ algebra but rather a commutative version of it, as $[(\widehat{\eta};0),(\widehat{\eta}';0)] = 0$. It is straightforward to show that

$$\pounds_{X_{\widehat{\xi}}}\widehat{\eta} = -\mathrm{ad}_{\widehat{\xi}}\widehat{\eta} = [\widehat{\eta},\widehat{\xi}] \tag{10.59}$$

so the product rule on $\widetilde{\mathfrak{g}}$ reads

$$[(\widehat{\eta};\widehat{\xi}),(\widehat{\eta}';\widehat{\xi}')] = ([\widehat{\eta},\widehat{\xi}'] - [\widehat{\eta}',\widehat{\xi}];[\widehat{\xi},\widehat{\xi}']) \tag{10.60}$$

where the commutators appearing on the right-hand side refer to the product on $\mathfrak{vira}$. Note that the elements $(\widehat{\eta};0)$, generating momentum translations, form an abelian (normal) subalgebra, and the elements $(0;\widehat{\xi})$, generating configuration translations, form a Virasoro subalgebra.

It is convenient to express this algebra in the usual *harmonic basis*, where the functions $\xi(\theta)$ and $\eta(\theta)$ are expanded in their Fourier modes. That is, we define the basis as

$$\begin{aligned} L_n &= (0; e^{in\theta}\partial_\theta)\,, \quad R = (0; \widehat{c}) \\ K_n &= (e^{in\theta}\partial_\theta; 0)\,, \quad T = (\widehat{c}; 0) \end{aligned} \tag{10.61}$$

where $n \in \mathbb{Z}$. In this basis, the algebra product reads

$$\begin{aligned} &[L_n, L_m] = i(n-m)L_{n+m} - 4\pi i n^3 \delta_{n+m,0} R \\ &[K_n, L_m] = i(n-m)K_{n+m} - 4\pi i n^3 \delta_{n+m,0} T \\ &[K_n, K_m] = 0 \\ &[R, \#] = 0 \\ &[T, \#] = 0 \end{aligned} \tag{10.62}$$

where δ is the Kronecker delta and "#" denotes "any element", so R and T form the center of the algebra.

10.3 Lifted group action on pre-phase spaces

We have described how the canonical group $\widetilde{G} = (\mathfrak{vira}^*)^* \rtimes \mathit{Vira}$ acts on the reduced phase space $\widetilde{\mathcal{P}} = T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$, but it is also useful to understand how this group action can be lifted to an action on the partially-reduced phase space $\widehat{\mathcal{S}} = \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$ and, ultimately, on the original (constrained) phase space

$\mathcal{P}$ coordinatized by ADM variables.

The action on $\widetilde{\mathcal{P}}$ by the group element $(\widehat{\eta};\widehat{\phi})$, denoted by $\Gamma_{(\widehat{\eta},\widehat{\phi})}$ and defined in (10.56), can be "pulled-back" or lifted to an action on $\widehat{\mathcal{S}}$, denoted by $J^*\Gamma_{(\widehat{\eta};\widehat{\phi})}$ and defined in such a way that the following diagram is commutative

$$\begin{array}{ccc} \widehat{\mathcal{S}} & \xrightarrow{J^*\Gamma} & \widehat{\mathcal{S}} \\ {\scriptstyle J}\big\downarrow & & \big\downarrow{\scriptstyle J} \\ \widetilde{\mathcal{P}} & \xrightarrow{\Gamma} & \widetilde{\mathcal{P}} \end{array}$$

that is, $J \circ (J^*\Gamma) = \Gamma \circ J$, where the subscript label $(\widehat{\eta};\widehat{\phi})$ of Γ and $J^*\Gamma$ has been omitted. Of course, being a group action, the map $(\widehat{\eta};\widehat{\phi}) \mapsto J^*\Gamma_{(\widehat{\eta};\widehat{\phi})}$ is required to be a homomorphism from $\widetilde{G}$ to the group of diffeomorphisms on $\widehat{\mathcal{S}}$.

In the definition of Γ, (10.55), let $p = J(\psi,\mathring{\sigma})$ so

$$\Gamma_{(\widehat{\eta},\widehat{\phi})} J(\psi,\mathring{\sigma}) = \mathrm{coad}^*_{\phi^{-1}} J(\psi,\mathring{\sigma}) - \widehat{\eta} \tag{10.63}$$

Note that it maps the 1-form $J(\psi,\mathring{\sigma})$ at the point $[\psi] = q(\psi)$ to a 1-form at $\phi[\psi] = \mathrm{coad}_\phi q(\psi)$, which is clear from the pull-back coadjoint map (also recall that $\widehat{\eta}$, seen as an element of $(\mathfrak{vira}^*)^*$, defines a 1-form at $\phi[\psi]$ by restricting it to the configuration space $\mathcal{Q}$, seen as the coadjoint orbit of $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$). Now let $\widetilde{X}$ be a vector tangent to $\mathcal{Q}$ at $[\phi\psi]$ defined as

$$\widetilde{X} := q_* l_{\phi\psi *}\xi \tag{10.64}$$

where $\xi \in \mathfrak{diff}(S^1)$. The whole tangent space at $[\phi\psi]$ is spanned by these vectors.

Applying the 1-form $\Gamma_{(\widehat{\eta},\widehat{\phi})}J(\psi,\mathring{\sigma})$ to $\widetilde{X}$ gives

$$\Gamma_{(\widehat{\eta},\widehat{\phi})}J(\psi,\mathring{\sigma})\widetilde{X} = J(\psi,\mathring{\sigma})(\text{coad}_{\phi^{-1}*}q_* l_{\phi\psi*}\xi) - \widehat{\eta}(\widetilde{X}) \tag{10.65}$$

Let us focus on the first term on the right-hand side. An identity that will be useful is $q \circ l_\phi = \text{coad}_\phi \circ q$ which can be shown as follows

$$q \circ l_\phi(\psi) = q(\phi\psi) = \text{coad}_{\phi\psi}\widetilde{\varepsilon} = \text{coad}_\phi \text{coad}_\psi \widetilde{\varepsilon} = \text{coad}_\phi \circ q(\psi) \tag{10.66}$$

where we have used formula (10.52) for $q : \mathit{Diff}^+(S^1) \to \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}) \subset \mathfrak{vira}^*$. The derivative of this identity allows us to get

$$J(\psi,\mathring{\sigma})(\text{coad}_{\phi^{-1}*}q_* l_{\phi\psi*}\xi) = J(\psi,\mathring{\sigma})(q_* l_{\phi^{-1}*} l_{\phi\psi*}\xi) = J(\psi,\mathring{\sigma})(q_* l_{\psi*}\xi) = \mathring{\sigma}(\xi) \tag{10.67}$$

where the definition of J has been used in the last step. Now let us consider the second term, $\widehat{\eta}(\widetilde{X})$. Define the notation $\text{coad}\,\widetilde{\varepsilon}(\psi) := \text{coad}_\psi\widetilde{\varepsilon}$, so that $q = \text{coad}\,\widetilde{\varepsilon}$. Note that, at the identity $\psi = I$, $q_* = (\text{coad}\,\widetilde{\varepsilon})_*$ is precisely equal to the coadjoint action of $\mathfrak{vira}$ on $\mathfrak{vira}^*$, i.e., $(\text{coad}\,\widetilde{\varepsilon})_*\xi = \text{coad}_\xi\widetilde{\varepsilon}$. Then

$$\widehat{\eta}(\widetilde{X}) = \widehat{\eta}(q_* l_{\phi\psi*}\xi) = \widehat{\eta}(\text{coad}_{\phi\psi*}q_*\xi) = \text{coad}^*_{\phi\psi}\widehat{\eta}(\text{coad}_\xi\widetilde{\varepsilon}) \tag{10.68}$$

As $\text{coad}^*_{\phi\psi}$ is a linear map from $\mathfrak{vira}$ into itself, the same argument used to go from

(10.55) to (10.56) applied here to replace $\text{coad}^*_{\phi\psi}$ by $\text{ad}_{(\phi\psi)^{-1}}$, yielding

$$\widehat{\eta}(\widetilde{X}) = \text{ad}_{(\phi\psi)^{-1}}\widehat{\eta}(\text{coad}_\xi\widetilde{\varepsilon}) \tag{10.69}$$

We then define the map $\Upsilon : \mathit{Diff}^+(S^1) \times \mathfrak{vira} \to \mathfrak{diff}^*(S^1)$ by

$$\Upsilon_\psi\widehat{\eta}(\xi) := \text{ad}_{\psi^{-1}}\widehat{\eta}(\text{coad}_\xi\widetilde{\varepsilon}) \tag{10.70}$$

As the notation suggests, it is convenient to think of Υ as $\mathit{Diff}^+(S^1)$-labeled linear map from $\mathfrak{vira}$ to $\mathfrak{diff}^*(S^1)$. We then have,

$$\Gamma_{(\widehat{\eta};\widehat{\phi})}J(\psi,\mathring{\sigma})\widetilde{X} = (\mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta})(\xi) \tag{10.71}$$

We need to infer what $J^*\Gamma$ has to be in order for $J \circ J^*\Gamma_{(\widehat{\eta};\widehat{\phi})}(\psi,\mathring{\sigma})\widetilde{X}$ to reproduce the same result. We know that $J^*\Gamma_{(\widehat{\eta};\widehat{\phi})}(\psi,\mathring{\sigma})$ must project to a 1-form at $[\phi\psi]$ under J, which suggests the obvious guess

$$J^*\Gamma_{(\widehat{\eta};\widehat{\phi})}(\psi,\mathring{\sigma}) = (\phi\psi, \mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta}) \tag{10.72}$$

This can be easily confirmed by projecting it under J and acting on $\widetilde{X}$, $J(\phi\psi,\mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta})(\widetilde{X}) = (\mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta})(\Xi(l_{\phi\psi*}\xi)) = (\mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta})(\xi)$, matching with (10.71). Note that the image of Υ is actually contained in the $\mathring{\mathfrak{diff}}^*(S^1)$ subspace of $\mathfrak{diff}^*(S^1)$, since every $\xi \in \mathfrak{psl}(2,\mathbb{R})$ is in the little group of $\widetilde{\varepsilon}$, so $\text{coad}_\xi\widetilde{\varepsilon} = 0$. Thus, for every

$\psi \in Diff^+(S^1)$, $\Upsilon_\psi : \mathfrak{vira} \to \mathring{\mathfrak{diff}}^*(S^1)$, and it makes sense to have $\Upsilon_{\phi\psi}\widehat{\eta}$ subtracted from $\mathring{\sigma}$ as it gives another element of $\mathring{\mathfrak{diff}}^*(S^1)$.

It is convenient to have a more explicit expression for $\Upsilon_\psi\widehat{\eta}$, since its definition (10.70) is quite abstract. Using expression (10.30) for $\mathrm{ad}_{\psi^{-1}}\widehat{\eta}$ and (10.39) for $\mathrm{coad}_\xi\widetilde{\varepsilon}$ we get

$$\Upsilon_\psi\widehat{\eta}(\xi) = -\int d\theta \frac{\eta(\psi(\theta))}{\psi'(\theta)}\left(2\xi'(\theta) + 2\xi'''(\theta)\right) = 2\int d\theta \left[\left(\frac{\eta(\psi(\theta))}{\psi'(\theta)}\right)' + \left(\frac{\eta(\psi(\theta))}{\psi'(\theta)}\right)'''\right]\xi(\theta) \tag{10.73}$$

where integration by parts was used in the last step. We have also used that $\psi_*^{-1}(\eta\partial_\theta)|_\theta = \frac{\eta(\psi(\theta))}{\psi'(\theta)}\partial_\theta$. Note that the central component of $\mathrm{ad}_{\psi^{-1}}\widehat{\eta}$ does not contribute since $\mathrm{coad}_\xi\widetilde{\varepsilon}$ has no central component. Therefore we can read off

$$\Upsilon_\psi\widehat{\eta} = 2\left[\left(\frac{\eta(\psi(\theta))}{\psi'(\theta)}\right)' + \left(\frac{\eta(\psi(\theta))}{\psi'(\theta)}\right)'''\right] d\theta^2 \tag{10.74}$$

which is manifestly belongs to $\mathring{\mathfrak{diff}}^*(S^1)$.

Let us briefly comment on how the group action lifts to the original (constrained) phase space, described in terms of metrics and extrinsic curvatures. The state $(\psi, \mathring{\sigma})$ corresponds to the metric $h_{ab} = \Psi_* e^\lambda \bar{h}_{ab}$ and the (traceless part of the) extrinsic curvature $\sigma^{ab} = \Psi_* e^{-2\lambda}\bar{\sigma}^{ab}$, where Ψ is any extension of ψ to the disc and λ satisfies the associated Lichnerowicz equation. Under a transformation $(\psi, \mathring{\sigma}) \mapsto (\psi', \mathring{\sigma}')$ there will be corresponding transformations of *(i)* $\Psi \mapsto \Psi'$, *(ii)* $\bar{\sigma}^{ab} \mapsto \bar{\sigma}'^{ab}$ and also of the Weyl scalar *(iii)* $\lambda \mapsto \lambda'$. Let us discuss each one of these changes under

$(\psi, \mathring{\sigma}) \mapsto J^*\Gamma_{(\widehat{\eta},\widehat{\phi})}(\psi, \mathring{\sigma}) = (\phi\psi, \mathring{\sigma} - \Upsilon_{\phi\psi}\widehat{\eta}) =: (\psi', \mathring{\sigma}')$:

(i) If Φ is any extension of ϕ to the disc, then $\Phi\Psi$ is an extension of $\phi\psi$ to the disc. We can therefore take $\Psi' = \Phi\Psi$.

(ii) Recall that $\mathring{\sigma}$ and $\bar{\sigma}^{ab}$ are related via

$$\mathring{\sigma}(\widehat{\xi}) = -\int \vartheta_{\bar{h}}\bar{\sigma}^{ab}\pounds_{\xi}\bar{h}_{ab} \tag{10.75}$$

Since $\Upsilon_{\phi\psi}\widehat{\eta} \in \mathring{\mathfrak{diff}}^*(S^1)$, we can use this same expression to define a symmetric traceless and divergenceless tensor $\bar{\Upsilon}^{ab}_{\phi\psi}\widehat{\eta}$, that is, $\Upsilon_{\phi\psi}\widehat{\eta}(\widehat{\xi}) = -\int\vartheta_{\bar{h}}\bar{\Upsilon}^{ab}_{\phi\psi}\widehat{\eta}\pounds_{\xi}\bar{h}_{ab}$. Then $\bar{\sigma}'^{ab} = \bar{\sigma}^{ab} - \bar{\Upsilon}^{ab}_{\phi\psi}\widehat{\eta}$.

(iii) The Weyl scalar λ changes for two reasons. First the "source term" $\bar{\sigma}^{ab}\bar{\sigma}_{ab}$ in the Lichnerowicz equation changes, and second the boundary values λ_∂ for that equation also changes. Note that λ_∂ satisfies $e^{\lambda|_\partial}d\theta^2 = \psi_*^{-1}\gamma$, so as $\psi \mapsto \phi\psi$ we have $e^{\lambda'_\partial}d\theta^2 = (\phi\psi)_*^{-1}\gamma$. While there is no explicit formula for λ', as it is a solution of this modified Lichnerowicz equation with modified boundary conditions, we know that the map $\lambda \mapsto \lambda'$ is well-defined due to the existence and uniqueness properties of solutions of that equation.

A very simple example to discuss is the $SO(2)$ rotation, i.e., the action of the group element $(\widehat{\eta}, \widehat{\phi}) = (0, r_\varphi)$ where $r_\varphi(\theta) = \theta + \varphi$ (modulo 2π). In that case the pre-phase space point changes as $(\psi, \mathring{\sigma}) \mapsto (r_\vartheta\psi, \mathring{\sigma})$. Note that $\bar{\sigma}^{ab}$ does not change. Also, λ_∂ does not change (using the convention that $\gamma \propto d\theta^2$) since $e^{\lambda'_\partial}d\theta^2 =$

$\psi_*^{-1}(R_\varphi)_*^{-1}\gamma = \psi_*^{-1}\gamma = e^{\lambda_\partial}d\theta^2$. Consequently the Weyl scalar λ does not change. Therefore, if R_φ is any extension of r_φ to the disc, $h'_{ab} = (R_\varphi\Psi)_* e^{\lambda}\bar{h}_{ab} = (R_\varphi)_* h_{ab}$ and $\sigma'^{ab} = (R_\varphi\Psi)_* e^{-2\lambda}\bar{\sigma}h_{ab} = (R_\varphi)_*\sigma^{ab}$, showing that $(\widehat{\eta},\widehat{\phi}) = (0, r_\varphi)$ really corresponds to a rotation of the initial data on the Cauchy slice.

11 The canonical charges for the diamond

In this subsection we shall compute the charges associated with the canonical group, $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$, and the corresponding Poisson algebra of them. As in (9.27), the momentum charges are associated with elements $(0;\widehat{\xi})$ and given by

$$P_{\widehat{\xi}}(p) = p(X_{\widehat{\xi}}) \tag{11.1}$$

where $p \in T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$ and $X_{\widehat{\xi}}$ is the vector field on $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, evaluated at $[\psi] = \pi(p)$, induced by $\widehat{\xi} \in \mathfrak{vira}$. And, as in (9.29), the configuration charges are associated with elements $(\widetilde{\eta};0)$ and given by

$$Q_{\widehat{\eta}}(p) = \widehat{\eta}(\pi(p)) \tag{11.2}$$

where $\pi(p)$ is seen as a vector in $\mathfrak{vira}^*$, as in (10.52).

Now we show that this choice of charges, with no constants added, form a legitimate representation of $\widetilde{\mathfrak{g}}$, i.e., without additional central charges.[50] In particular,

[50] The reason for saying "additional" central charges is because $\widetilde{\mathfrak{g}}$ already has a non-trivial center given by $(\widehat{c};0) \oplus (0;\widehat{c})$.

this implies that there is no need to centrally extend the group $\widetilde{G}$ by a 2-cocycle. As explained in subsection 9.3, the only 2-cocycle that can be non-trivial is the one mixing the P and Q parts of the algebra, i.e., $z((\widehat{\eta},0),(0,\widehat{\xi}))$. According to (9.42),

$$\{Q_{\widehat{\eta}}, P_{\widehat{\xi}}\} = \widehat{\eta}(X_{\widehat{\xi}}) \tag{11.3}$$

Since z is constant on the phase space, it suffices to evaluated it at a single point. Most conveniently, we choose to evaluate it at $p = p_0 = 0$ on the fiber over $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$. At this point, $X_{\widehat{\xi}} = \mathrm{coad}_{\widehat{\xi}}\widetilde{\varepsilon}$, so we have

$$\{Q_{\widehat{\eta}}, P_{\widehat{\xi}}\}(p_0) = \widehat{\eta}(\mathrm{coad}_{\widehat{\xi}}\widetilde{\varepsilon}) = \mathrm{coad}_{\widehat{\xi}}\widetilde{\varepsilon}(\widehat{\eta}) = -\widetilde{\varepsilon}(\mathrm{ad}_{\widehat{\xi}}\widehat{\eta}) = \widetilde{\varepsilon}([\widehat{\eta},\widehat{\xi}]) \tag{11.4}$$

But from expression (11.2) for the charge, and from the product rule (10.60), we have

$$H_{[(\widehat{\eta};0),(0;\widehat{\xi})]} = H_{([\widehat{\eta},\widehat{\xi}];0)} = Q_{[\widehat{\eta},\widehat{\xi}]} \tag{11.5}$$

which evaluated at p_0 gives

$$H_{[(\widehat{\eta};0),(0;\widehat{\xi})]}(p_0) = [\widehat{\eta},\widehat{\xi}](\widetilde{\varepsilon}) = \widetilde{\varepsilon}([\widehat{\eta},\widehat{\xi}]) \tag{11.6}$$

matching exactly with $\{Q_{\widehat{\eta}}, P_{\widehat{\xi}}\}$. Therefore $z((\widehat{\eta};0),(0;\widehat{\xi})) = 0$.

The harmonic basis in (10.61) defines a basis of charges given by

$$P_n := H_{L_n}$$
$$Q_n := H_{K_n} \tag{11.7}$$

together with the central charges H_R and H_T. From the general expressions for the charges, we can compute explicitly the values of the central charges (which are constants on the phase space). We get, at a point $p \in T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$,

$$H_R = p(X_{\widehat{c}}) = 0 \tag{11.8}$$

where $X_{\widehat{c}} = 0$ because the coadjoint action of $\widehat{c}$ is trivial. If ψ is such that $\pi(p) = \mathrm{coad}_\psi \widetilde{\varepsilon}$, then

$$H_T = \widehat{c}(\pi(p)) = \widehat{c}(\mathrm{coad}_\psi \widetilde{\varepsilon}) = 1 \tag{11.9}$$

where we have used that the $\widetilde{c}$ component of $\mathrm{coad}_\psi \widetilde{\varepsilon}$ is always $1\widetilde{c}$, for any ψ, as in (10.52). In this way, the Poisson algebra of the canonical charges becomes

$$\{P_n, P_m\} = i(n-m)P_{n+m}$$
$$\{Q_n, P_m\} = i(n-m)Q_{n+m} - 4\pi i n^3 \delta_{n+m,0}$$
$$\{Q_n, Q_m\} = 0 \tag{11.10}$$

revealing that the algebra of momentum charges is just isomorphic to $\mathfrak{diff}(S^1)$, after all. That is, the whole effect of having taken the Virasoro group as the group of "configuration translations", was not really to extend the algebra of configuration translations, but rather to introduce a central charge in the commutators mixing the configuration and momentum charges. It turns out that this algebra is equal to a $\mathfrak{bms}_3$ algebra of asymptotic symmetries of asymptotically flat three-dimensional gravity, as described in [43, 45].

There are particularly explicit expressions for the charges on the partially-reduced phase space $\widehat{\mathcal{S}} = \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$. We start by pulling back the P charges under the J map. If $p = J(\psi, \mathring{\sigma})$, we have

$$P_{\widehat{\xi}}(p) = J(\psi, \mathring{\sigma})(X_{\widehat{\xi}}) = \mathring{\sigma}(l^*_{\psi^{-1}} \overline{X}_{\widehat{\xi}}) \tag{11.11}$$

where $\overline{X}_{\widehat{\xi}}$ is any vector at ψ that projects to $X_{\widehat{\xi}}$ under q. Given $\widetilde{\alpha} \in \mathfrak{vira}^*$, define the coadjoint map from Vira to $\mathfrak{vira}^*$ as

$$\operatorname{coad} \widetilde{\alpha}(\widehat{\psi}) := \operatorname{coad}_{\widehat{\psi}} \widetilde{\alpha} \tag{11.12}$$

Naturally, since the central elements of Vira have a trivial coadjoint action, we can also think of the same map as being from $\mathit{Diff}^+(S^1)$ to $\mathfrak{vira}^*$ by simply replacing $\widehat{\psi}$ by ψ. Note that derivative of this map gives the coadjoint action of $\mathfrak{vira}$ on $\mathfrak{vira}^*$,

$$(\operatorname{coad} \widetilde{\alpha})_* \widehat{\xi} = \operatorname{coad}_{\widehat{\xi}} \widetilde{\alpha} \tag{11.13}$$

In this notation, the quotient map $q : \mathit{Diff}^{+}(S^1) \to \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$, as realized in (10.52), is simply written as

$$q = \mathrm{coad}\,\widetilde{\varepsilon} \tag{11.14}$$

where $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$, and the vector field X_ξ induced by $\xi \in \mathfrak{diff}(S^1) \subset \mathfrak{vira}$ at $\widetilde{\alpha} \in \mathfrak{vira}^*$ is given by

$$X_\xi = (\mathrm{coad}\,\widetilde{\alpha})_* \xi \tag{11.15}$$

If $\psi \in \mathit{Diff}^{+}(S^1)$ is such that $\widetilde{\alpha} = \mathrm{coad}_\psi \widetilde{\varepsilon}$, we get

$$X_\xi = \left(\mathrm{coad}(\mathrm{coad}_\psi \widetilde{\varepsilon})\right)_* \xi \tag{11.16}$$

Observe that, given any $\phi \in \mathit{Diff}^{+}(S^1)$, we have

$$\mathrm{coad}(\mathrm{coad}_\psi \widetilde{\varepsilon})(\phi) = \mathrm{coad}_\phi \mathrm{coad}_\psi \widetilde{\varepsilon} = \mathrm{coad}_{\phi\psi} \widetilde{\varepsilon} = \mathrm{coad}\,\widetilde{\varepsilon}(\phi\psi) = \mathrm{coad}\,\widetilde{\varepsilon} \circ r_\psi(\phi) \tag{11.17}$$

where r_ψ is the right-multiplication in $\mathit{Diff}^{+}(S^1)$. Thus,

$$X_\xi = (\mathrm{coad}\,\widetilde{\varepsilon})_* r_{\psi *} \xi = q_* r_{\psi *} \xi = q_* l_{\psi *} l_{\psi^{-1} *} r_{\psi *} \xi = q_* l_{\psi *} (\mathrm{ad}^{\mathfrak{diff}}_{\psi^{-1}} \xi) \tag{11.18}$$

where l_ψ is the left-multiplication in $\mathit{Diff}^{+}(S^1)$ and $\mathrm{ad}^{\mathfrak{diff}}$ is the adjoint map in $\mathfrak{diff}(S^1)$. This reveals that $\overline{X}_{\widehat{\xi}}$ can be chosen as $l_{\psi *}(\mathrm{ad}^{\mathfrak{diff}}_{\psi^{-1}} \xi)$, and therefore

$$P_{\widehat{\xi}} = J(\psi, \mathring{\sigma})(X_{\widehat{\xi}}) = \mathring{\sigma}(\mathrm{ad}^{\mathfrak{diff}}_{\psi^{-1}} \xi) = \mathrm{coad}^{\mathfrak{diff}}_{\psi} \mathring{\sigma}(\xi) \tag{11.19}$$

If $\mathring{\sigma} \in \mathring{\mathfrak{diff}}^*(S^1)$ is seen as an element of $\mathfrak{vira}^*$, $\mathring{\sigma} = \mathring{\sigma} + 0\widetilde{c}$, we have that $\mathrm{coad}_\psi^{\mathfrak{diff}} \mathring{\sigma} = \mathrm{coad}_\psi \mathring{\sigma}$ (the two coadjoint maps always match when acting on elements with no $\widetilde{c}$ components). So we can write

$$P_{\widehat{\xi}} = \mathrm{coad}_\psi \mathring{\sigma}(\xi) = (\psi_* \mathring{\sigma})(\xi) \tag{11.20}$$

or, a little bit more explicitly,

$$P_{\widehat{\xi}}(\psi, \mathring{\sigma}) = \int d\theta \, \frac{\mathring{\sigma}(\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} \xi(\theta) \tag{11.21}$$

Now we pull-back the Q charges under J, which is trivial since it is enough to write $\pi(p) = \mathrm{coad}_\psi \widetilde{\varepsilon}$, yielding

$$Q_{\widehat{\eta}}(p) = \widehat{\eta}(\mathrm{coad}_\psi \widetilde{\varepsilon}) \tag{11.22}$$

or, more explicitly,

$$Q_{\widehat{\eta}}(\psi, \mathring{\sigma}) = \int d\theta \, \frac{1 - 2S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} \eta(\theta) + y \tag{11.23}$$

where $\widehat{\eta} = \eta(\theta)\partial_\theta + y\widehat{c}$.

11.1 Geometrical interpretation of the charges

Statements made about a quantum theory are only useful to the extent that one can interpret their classical consequences. That is, in order to extract physical information from a quantum theory one needs not only to know the properties of some self-adjoint operators, but it is important that one can match those operators to classical observables, which in principle could be tested and understood through classical experiments. So far we have already encountered one observable to which we could give a simple geometric interpretation: the CMC time-evolution Hamiltonian which corresponds, at a time τ, to the area of the CMC slice with $K = -\tau$. In the previous section the canonical charges were defined quite abstractly on the reduced phase space $\widetilde{\mathcal{P}}$, and were also given in terms of $(\psi, \mathring{\sigma}) \in \widehat{\mathcal{S}}$, which are still abstract parameters characterizing the diamond. In order to understand their physical meaning we should try to express them in terms of something more concrete like ADM variables, (h_{ab}, K^{ab}), for these have a simple geometrical meaning (the metric and extrinsic curvature of the CMC slices).

Note: This section relies heavily on the content and notation of Part I. Importantly, one should notice that in this section $\Psi \in \mathit{Diff}^+(\Sigma)$ denotes an extension of $\psi \in \mathit{Diff}^+(S^1)$ to the Cauchy slice, and should not be confused with quantum wavefunctions that will appear later.

11.1.1 Momentum charges

Let us begin with the P charges. Formula (11.20), in conjunction with (10.75), gives

$$P_\xi = \mathring{\sigma}(\psi_*^{-1}\xi) = -\int \vartheta_{\bar{h}} \bar{\sigma}^{ab} \pounds_{\Psi_*^{-1}\xi} \bar{h}_{ab} \tag{11.24}$$

where in the last expression Ψ is any extension of ψ to a diffeomorphism of the disk and ξ is, with a slight abuse of notation, any smooth extension of ξ (at the boundary) to a vector field on the disk. Now we insert factors of e^{λ}, where λ is the solution of the associated Lichnerowicz equation, as follows

$$P_{\xi} = \mathring{\sigma}(\psi_*^{-1}\xi) = -\int e^{\lambda}\vartheta_{\bar{h}} e^{-2\lambda}\bar{\sigma}^{ab} e^{\lambda} \pounds_{\Psi_*^{-1}\xi}\bar{h}_{ab} \tag{11.25}$$

Using relation $\vartheta_{e^{\lambda}\bar{h}} = e^{\lambda}\vartheta_{\bar{h}}$ and $h_{ab} = \Psi_* e^{\lambda}\bar{h}_{ab}$ we have $e^{\lambda}\vartheta_{\bar{h}} = \vartheta_{\Psi_*^{-1}h}$. Also, notice that $e^{\lambda}\pounds_{\Psi_*^{-1}\xi}\bar{h}_{ab} = \pounds_{\Psi_*^{-1}\xi}\left(e^{\lambda}\bar{h}_{ab}\right) - (\pounds_{\Psi_*^{-1}\xi}e^{\lambda})\bar{h}_{ab}$, but the second term would vanish when contracted with $\bar{\sigma}^{ab}$, so we can replace $e^{\lambda}\pounds_{\Psi_*^{-1}\xi}\bar{h}_{ab} = \pounds_{\Psi_*^{-1}\xi}\left(e^{\lambda}\bar{h}_{ab}\right)$ in the integrand. Finally, from $e^{\lambda}\bar{h}_{ab} = \Psi_*^{-1}h_{ab}$ and $\sigma^{ab} = \Psi_* e^{-2\lambda}\bar{\sigma}^{ab}$ we have $e^{-2\lambda}\bar{\sigma}^{ab} = \Psi_*^{-1}\sigma^{ab}$. Thus, we get

$$P_{\xi} = -\int \vartheta_{\Psi_*^{-1}h}\Psi_*^{-1}\sigma^{ab}\pounds_{\Psi_*^{-1}\xi}\left(\Psi_*^{-1}h_{ab}\right) = -\int \Psi_*^{-1}\left[\vartheta_h \sigma^{ab}\pounds_{\xi}h_{ab}\right] \tag{11.26}$$

and since integrals are invariant under diffeomorphisms,

$$P_{\xi} = -\int \vartheta_h \sigma^{ab}\pounds_{\xi}h_{ab} \tag{11.27}$$

As σ^{ab} is transverse with respect to h_{ab}, this can be converted to a boundary integral

$$P_{\xi} = -2\int \vartheta_h \sigma^{ab}\nabla_a\xi_b = -2\int_{\partial} d\theta\, \sigma^{ab} n_a \xi_b \tag{11.28}$$

where n^a is the unit (outward-pointing) vector field normal to the boundary. Finally, using $\sigma^{ab} := K^{ab} - \frac{1}{2}K h^{ab}$ and the orthogonality between n and ξ (at the boundary),

$$P_\xi = -2 \int_\partial d\theta \, K_{ab} n^a \xi^b \tag{11.29}$$

which is fully expressed in terms of geometrical quantities.

Since we are interested in the physical interpretation, it is worthwhile to reintroduce the physical scales, as explained in Sec. 8.1. We note that (11.24) remains unchanged, but there are factors of ℓ and ℓ_P in the transformation to h_{ab} and σ^{ab}, so (11.28) becomes

$$P_\xi = -\frac{1}{8\pi \ell_P} \int_\partial ds \, K_{ab} n^a \xi^b \tag{11.30}$$

where ds is the element of length with respect to h_{ab}. Now we wish to re-express ξ, which is an abstract vector on S^1, in terms of a vector associated with the physical boundary of the Cauchy slice. If t^a is the unit vector tangent to the boundary of the diamond, $h_{ab} t^a t^b = 1$, and if we assume that the physical boundary metric is[51]

$$\gamma = \left(\frac{\ell}{2\pi} \right)^2 d\theta^2 \tag{11.31}$$

then

$$t^a = \frac{2\pi}{\ell} (\partial_\theta)^a \tag{11.32}$$

[51] There seems to be no reason not to make this assumption, since any γ can be parametrized by proper length so that $\gamma = ds^2$, and then the angle coordinate θ, with respect to which $\xi = \xi(\theta)\partial_\theta$, can be defined simply by $\theta := 2\pi s/\ell$. Later, in Sec. 12.5, we will argue that different choices lead to physically equivalent theories, even at the quantum level.

Therefore, with this choice $\xi^a = \xi(\theta)(\partial_\theta)^a = \xi(\theta)(\ell/2\pi)t^a$, we have

$$P_\xi = -\frac{\ell}{16\pi^2 \ell_P} \int_\partial ds\, K_{ab} n^a \xi(\theta) t^b \tag{11.33}$$

where $\theta = 2\pi(s/\ell)$, with s being the length along the boundary from some reference point (a change of this reference point amounts to redefining the charges by phases, $P_n \mapsto e^{in\varphi} P_n$ and $Q_n \mapsto e^{in\varphi} Q_n$, which clearly leaves the canonical algebra invariant).

Let us discuss the P_0 charge in more depth, as it has some interesting features. From the expression above, with $\xi = \partial_\theta$, we see that

$$P_0 = -\frac{\ell}{16\pi^2 \ell_P} \int_\partial ds\, K_{ab} n^a t^b \tag{11.34}$$

Notably, P_0 generates a $SO(2)$ dynamical symmetry of the diamond, corresponding to rigid rotations of the boundary. It is therefore natural to call P_0 the *spin* of the diamond. This can be seen in two ways. First, as $P_0 = P_{\partial_\theta}$, its Poisson flow is generated by ∂_θ acting on the configuration space $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$. More precisely, a finite rotation by the angle φ, acting on S^1 as $R_\varphi(\theta) = \theta + \varphi$, will act on the configuration space as $R_\varphi[\psi] = [R_\varphi \circ \psi]$. We can show explicitly that the Hamiltonian is invariant under such a transformation, so that $\{P_0, \widetilde{H}\} = 0$, implying that P_0 is conserved under time evolution.

Another way to see that P_0 corresponds to a rotational symmetry is by considering directly the ADM formalism. In this context the Hamiltonians takes the form

$$H_{bulk}[N,\vec{N}] = \frac{1}{16\pi G}\int_\Sigma d^2x \left[N\sqrt{h}\left(K^{ab}K_{ab} - K^2 - R + 2\Lambda\right) - 2N_b\nabla_a\pi^{ab}\right] \tag{11.35}$$

up to boundary terms. The Hamiltonian, labelled by a lapse N and a shift $\vec{N}$ parameters, can be interpreted as generating spacetime diffeomorphisms along the vector $Nu^a + N^a$, where u^a is the unit future-directed vector normal to the spatial slice. When the lapse and shift are non-trivial at the boundary, the expression above may not define a regular Hamiltonian function on the phase space, i.e., a function that can be associated with a regular Hamiltonian flow. More precisely, note that this association is defined via $\delta H = -\imath_X\Omega$, so if the symplectic form Ω is given by the integral of a symplectic current on the Cauchy slice without boundary terms, $\Omega = \int_\Sigma \delta\pi^{ab}\wedge\delta h_{ab}$, then the right-hand side of this relation gives $\int_\Sigma (X^h)_{ab}\delta\pi^{ab} - (X^\pi)^{ab}\delta h_{ab}$, revealing that the left-hand side must be of the form $\delta H = \int_\Sigma A^{ab}\delta h_{ab} + B_{ab}\delta\pi^{ab}$ for a solution to exist. In other words, H generates a regular symplectic flow if and only if it has a well-defined functional derivative. In this case we say that H is *symplectically differentiable*. Since the Poisson bracket is defined in terms of the associated Hamiltonian vectors, functions only have well-defined Poisson brackets if they are symplectically differentiable. Accordingly, we may need to add appropriate boundary terms to expression (11.35) in order to define regular Hamiltonians. A trick to find these boundary terms is to vary the bulk

Hamiltonian H_{bulk} given in (11.35) and, if this produces boundary terms, we add suitable boundary terms H_∂ to cancel it, so that $\delta H = \delta H_{bulk} + \delta H_\partial$ is purely a bulk integral (not containing spatial derivatives of field variations). If the Hamiltonian does acquire a non-trivial boundary term in this manner, then its on-shell value will be $H \approx H_\partial$, implying that $\delta H = -\imath_X \Omega \neq 0$ and consequently the phase space transformation X induced by H is not gauge. In other words, the diffeomorphism generated by N and N^a would not be a gauge transformation but rather a non-trivial symmetry. Let us consider the case where N vanishes at the boundary and $N^a|_\partial$ is tangent to the boundary, i.e., $N^a \propto t^a$.[52] The variation of the term involving N in (11.35) gives

$$\delta H_{bulk} = \text{“}\textit{bulk term involving } N\text{”} - \frac{1}{16\pi\ell_P}\int_\partial ds\, N n^a \left(\nabla^b \delta h_{ab} - h^{cd}\nabla_a \delta h_{cd}\right) + \cdots \tag{11.36}$$

where n^a is the unit spatial outward-pointing normal vector field on the boundary, and the "$\cdots$" refers to terms involving N^a. Since we chose $N|_\partial = 0$, no boundary terms come from N. Now the term involving N^a gives

$$\delta H_{bulk} = \text{“}\textit{bulk terms}\text{”} - \frac{1}{8\pi\ell_P}\int_\partial ds\, n_a N_b \left(h^{-1/2}\delta\pi^{ab} + h^{bc}\delta h_{cd} h^{-1/2}\delta\pi^{ad}\right) \tag{11.37}$$

[52] A diffeomorphism generated by N^a would not in general preserve the boundary metric, unless $N^a \propto t^a$ by a constant factor. Nevertheless, in Sec. (11.1.3) we will argue that there is a way in which these Hamiltonians produce well-defined flows by virtue of the CMC gauge-fixing constraint.

This boundary term is in fact exact (as a phase space form) and can be written as

$$\delta H_{bulk} = \text{“}\textit{bulk terms}\text{”} - \delta\left[\frac{1}{8\pi\ell_P}\int_\partial ds\, n_a N_b K^{ab}\right] \tag{11.38}$$

which implies that the Hamiltonian must be "corrected" with a boundary term as follows

$$H = H_{bulk} + \frac{1}{8\pi\ell_P}\int_\partial ds\, n_a N_b K^{ab} \tag{11.39}$$

Thus, if N^a is non-zero at the boundary, the on-shell value of the corresponding Hamiltonian is non-trivial

$$H[N^a|_\partial = \xi^a]_{\textit{on-shell}} = \frac{1}{8\pi\ell_P}\int_\partial ds\, n_a \xi_b K^{ab} \tag{11.40}$$

Using convention (11.31) for the boundary metric, so that $\xi^a = \xi(\theta)(\partial_\theta)^a = \xi(\theta)(\ell/2\pi)t^a$, we get

$$H[N^a|_\partial = \xi^a]_{\textit{on-shell}} = \frac{\ell}{16\pi^2\ell_P}\int_\partial ds\, n_a \xi(\theta) t_b K^{ab} = -P_\xi \tag{11.41}$$

In particular, when N^a is an isometry of γ, $\xi = \partial_\theta$,

$$H[N^a|_\partial = (\partial_\theta)^a]_{\textit{on-shell}} = \frac{\ell}{16\pi^2\ell_P}\int_\partial ds\, n_a t_b K^{ab} = -P_0 \tag{11.42}$$

Thus we see that this Hamiltonian, generator of rigid rotations of the boundary, corresponds to (minus) the charge $P_0 = P_{\partial_\theta}$, justifying its denomination as the *spin* of the diamond. (The reason for the minus sign is due to our conventions of how

Diff$^+(S^1)$, or *Vira*, acts on the phase space.)

We can show that P_0 is preserved under CMC time evolution (or, equivalently, the Hamiltonian generating CMC time evolution is rotation-invariant). In fact, P_0 is independent of the spatial slice altogether. Let $H_{\vec{N}}$ be the Hamiltonian corresponding to a shift N^a that matches $(\partial_\theta)^a$ at the boundary and no lapse, which as we have just seen reduces to $-P_0$ on-shell; and let H_N be the Hamiltonian corresponding to a lapse N which deforms one Cauchy slice into another (thus vanishing at the boundary) and no shift. The change of P_0 with respect to slice change is then given by

$$\begin{aligned}
\dot{P}_0 = -\{H_{\vec{N}}, H_N\} &= -\{H_{\vec{N}}, \int d^2x\, N(x)\mathcal{H}_0(x)\} \\
&= -\int d^2x\, N(x)\{H_{\vec{N}}, \mathcal{H}_0(x)\} \\
&= \int d^2x\, N \pounds_{\vec{N}}\mathcal{H}_0 \\
&= \int d^2x\, N\left(\nabla_a N^a \mathcal{H}_0 + N^a \nabla_a \mathcal{H}_0\right) \\
&= \int d^2x\, \nabla_a\left(N N^a \mathcal{H}_0\right) - \int d^2x\, N^a \nabla_a N \mathcal{H}_0 \\
&= \int_\partial ds\, N n_a N^a (h^{1/2}\mathcal{H}_0) + H_{\pounds_{\vec{N}} N} \\
&= H_{\pounds_{\vec{N}} N} \approx 0
\end{aligned} \tag{11.43}$$

In the first line we wrote H_N explicitly in terms of the Hamiltonian constraint $\mathcal{H}_0$; in the third line we used that $H_{\vec{N}}$ generates spatial diffeomorphisms; in the fourth line we wrote the Lie derivative in terms of covariant derivatives (note that $\mathcal{H}_0$ is a density); in the fifth line we integrated by parts; in the sixth line we used Gauss' theorem and identified the bulk term as the Hamiltonian corresponding to the lapse $\pounds_{\vec{N}} N$; and in the last line we used that N^a is tangent to the boundary (so $n_a N^a = 0$) leading to the conclusion that $\dot{P}_0$ vanishes on-shell (i.e., on the constraint surface).

11.1.2 The spin/twist relationship

After having geometrically expressed P_0 as an integral of the extrinsic curvature along the boundary and recognized it as the spin of the diamond, we now provide another nice interpretation for it as being the *twist of the corner of the diamond*. First, let us define the twist of a curve embedded in a three-dimensional space. Consider a closed spacelike curve $q : S^1 \to \mathcal{M}$ embedded in an oriented Lorentzian space $\mathcal{M}$.[53] Let t be the unit vector tangent to q and $a^a := t^b \boldsymbol{\nabla}_b t^a$ be the acceleration vector. Since t is normalized, a is orthogonal to t, so let us define n as the unit vector aligned with a (assumed not be lightlike or zero). The plane tn defines a frame along the curve, whose third normal vector can be defined as $u^a := \boldsymbol{\epsilon}^{abc} t_b n_c$, where $\boldsymbol{\epsilon}$ is the volume element on $\mathcal{M}$. The *torsion* quantifies the rate (with respect to proper distance) that this frame rotates along the curve, and can be defined as $\chi := sign(n) u_a (t^b \boldsymbol{\nabla}_b n^a) = sign(n) \boldsymbol{\epsilon}_{abc} t^b n^c (t^d \boldsymbol{\nabla}_d n^a)$, where $sign(n) := n_a n^a$. The *twist* $\mathcal{T}$ of the (closed) curve is then defined as the integrated torsion along the curve,

$$\mathcal{T} := sign(n) \int ds \, \boldsymbol{\epsilon}_{abc} t^b n^c (t^d \boldsymbol{\nabla}_d n^a) \tag{11.44}$$

[53]The concepts discussed here also apply naturally for Riemmanian spaces.

where ds is the element of length along the curve.[54] What is interesting about the twist is that it can be evaluated with the expression above using *any* unit normal vector n, not necessarily the one aligned with the acceleration. This can be seen by evaluating this expression with two unit normal vectors n and $\widetilde{n}$. If both vectors are spacelike, we can write $\widetilde{n} = \cosh(\varphi)n + \sinh(\varphi)u$, where $\varphi : S^1 \to \mathbb{R}$ denotes the (hyperbolic) angle of rotation between n and $\widetilde{n}$. We get

$$\boldsymbol{\epsilon}_{abc} t^b \widetilde{n}^c (t^d \boldsymbol{\nabla}_d \widetilde{n}^a) = \boldsymbol{\epsilon}_{abc} t^b n^c (t^d \boldsymbol{\nabla}_d n^a) - \frac{d\varphi}{ds} \tag{11.45}$$

If both n and $\widetilde{n}$ are smooth along the curve, so that φ is periodic, the integral of $d\varphi/ds$ will vanish, and therefore the twist can be defined as in (11.44) for any (smooth) spacelike unit normal vector field n. One can verify that the definition also applies for timelike normal vectors, thanks to the factor of *sign*(n). Note that the twist is always well-defined, even when the torsion is not (the torsion is only well-defined if the acceleration is not vanishing and not lightlike).

We can also provide a formula for the twist which refers to lightlike normal vectors. This may be pertinent since the two inward-pointing null vectors are the generators of the future and past horizons of the diamond. Let $k_\pm$ be the future (+) and past (−) inward-pointing null vector fields orthogonal to the curve, normalized like $(k_+)_a(k_-)^a = 1$. If n is a smooth (outward-pointing) spacelike normal field along the curve, and u is the (future-pointing) vector field normal to both n and t,

[54]Some references include numerical factors like $1/2\pi$ in the definition of the twist, which we choose not to do.

then we can write $k_\pm = (-n \pm u)/\sqrt{2}$. Now writing $\mathcal{T} = (\mathcal{T} + \mathcal{T})/2$, where the first $\mathcal{T}$ is expressed as in (11.44) with n as the normal vector and the second $\mathcal{T}$ with u as the normal vector, we see that

$$\mathcal{T} = \int ds\, \boldsymbol{\epsilon}_{abc} t^b (k_\pm)^c t^d \boldsymbol{\nabla}_d (k_\mp)^a = \mp \int ds\, (k_\pm)_a t^b \boldsymbol{\nabla}_b (k_\mp)^a \tag{11.46}$$

Note that this involves both k_+ and k_-, essentially evaluating how much one changes with respect to the other.

Interestingly, relation (11.45) provides another interpretation for the twist, as follows. Let $\widetilde{n}$ be a unit (spacelike) normal Fermi-Walker transported[55] along the curve, thus satisfying

$$D_q^{\mathrm{FW}} \widetilde{n}^a = t^b \boldsymbol{\nabla}_b \widetilde{n}^a + a_b \widetilde{n}^b t^a = 0 \tag{11.47}$$

When $\widetilde{n}$ comes back to the initial point, after a loop, it may not coincide with its original value. If n is any smooth (spacelike) unit normal field along the curve, we can define the (hyperbolic) angle φ of $\widetilde{n}$ with respect to n as before. But now φ is not necessarily periodic, so let $\Delta\varphi := \varphi_{\text{final}} - \varphi_{\text{initial}}$ be the total angle accumulation after a loop. Integrating both sides of (11.45) in ds yields zero for the left-hand side (since $\boldsymbol{\epsilon}$ would be contracting with two t's), $\mathcal{T}$ for the first term on the right-hand

[55]The Fermi-Walker transport gives a natural notion of "parallel transporting" a frame along a curve, while keeping it aligned with the curve. More precisely, this is defined by parallel-transporting a frame for an infinitesimal distance and then projecting the normal vectors orthogonally to the curve and the tangent vector along the curve. The Fermi-Walker transport coincides with the parallel transport if (and only if) the curve is a geodesic. The Fermi-Walker derivative is defined by $D_q^{\mathrm{FW}} v^a := t^b \boldsymbol{\nabla}_b v^a + (a_b t^a - t_b a^a) v^b$, where t and a are respectively the tangent and acceleration vectors of the curve q, and it has the property that $D_q^{\mathrm{FW}} v^a = 0$ if (and only if) v is Fermi-Walker transported along the curve.

side and $-\Delta\varphi$ for the last term. Therefore, we have

$$\mathcal{T} = \Delta\varphi \tag{11.48}$$

revealing that the twist $\mathcal{T}$ can be interpreted as a holonomy induced on the normal bundle of the curve, expressing how the parallel-transported normal frame comes back rotated after a loop. A similar result also holds for lightlike vectors, using (11.46). In this case, let $\widetilde{k}_\pm$ be a pair of Fermi-Walker transported inward-pointing null vectors orthogonal to the curve, normalized like $(\widetilde{k}_+)_a(\widetilde{k}_-)^a = 1$, and let $k_\pm$ be any smooth (normalized) pair of inward-pointing null vectors orthogonal to the curve. We can write $\widetilde{k}_\pm = e^{\chi_\pm}k_\pm$ for some scalars $\chi_\pm$. The normalization conditions imply that $\chi_- = -\chi_+$. From a relation analogue to (11.45), for the integrand of (11.46), we conclude that

$$\mathcal{T} = \mp\Delta\chi_\pm \tag{11.49}$$

where $\Delta\chi_\pm := (\chi_\pm)_{\text{final}} - (\chi_\pm)_{\text{initial}}$ is the phase difference accumulated along a loop. That is, the future and past null generators of the Cauchy horizon of the diamond come back boosted (rescaled with respect to each other) after being transported along the corner loop of the diamond.

Let us now see that P_0 is directly related to the twist of the boundary. Let u be the unit timelike (future-pointing) normal vector to the CMC, so that $K_{ab} = \boldsymbol{\nabla}_b u_a$, where $\boldsymbol{\nabla}$ denotes the spacetime covariant derivative. Let n be the unit outward-pointing normal vector to the boundary, parallel to the CMC, and t^a be the unit vector tangent to the boundary, assumed to be oriented such that $\boldsymbol{\epsilon}(u, n, t) = 1$, where $\boldsymbol{\epsilon}$ is the spacetime orientation. Then (11.34) can be integrated by parts as

$$P_0 = -\frac{\ell}{16\pi^2 \ell_P} \int_\partial ds\, \boldsymbol{\nabla}_b u_a n^a t^b = \frac{\ell}{16\pi^2 \ell_P} \int_\partial ds\, u_a t^b \boldsymbol{\nabla}_b n^a = \frac{\ell}{16\pi^2 \ell_P} \mathcal{T} \qquad (11.50)$$

showing that the spin P_0 is proportional to the twist $\mathcal{T}$, with a proportionality factor involving the ratio between the boundary length and the Planck length (in units where $\hbar = 1$).

This relationship between twist and spin can be seen as an analogue, in lower dimensions, of a result obtained in [102], in which the edge mode structure of gravity in 3+1 dimensions is analyzed from the covariant phase space perspective. It is found that the charges generating volume-preserving diffeomorphisms of the corner (S^2) are related to the curvature of the connection on the normal bundle of the corner, as naturally defined from its embedding into the ambient spacetime. For our causal diamonds, volume-preserving diffeomorphisms of the corner (S^1) are precisely the isometries of the boundary metric. In this case, the normal bundle of the corner also inherits a natural connection from the ambient spacetime, but one cannot define a non-trivial curvature tensor since the base space is 1-dimensional. Nevertheless, due to the non-trivial fundamental group of S^1 (unlike S^2), there is another kind of invariant associated to the connection: the holonomy defined by parallel transporting the normal frame along the loop. As we have seen, this holonomy is described by a boost angle which is interpreted as the twist of the corner.

On a separate note, the twist of a curve has also appeared in an analysis of the holographic entanglement entropy in the context of two-dimensional conformal field theories with a gravitational anomaly [103], but it is unclear whether this has any relation with our work.

11.1.3 Configuration charges

The interpretation of the Q charges appears to be much less simple. Unlike the P charges which are related to a local integration of K_{ab} along the boundary, the Q charges seem to be related in a non-local way with the curvature of the boundary as embedded in the CMC itself.

We can still establish some general properties of the Q charges. First, note from (11.23) that the Q's do not depend on $\mathring{\sigma}$, but only on ψ. In fact, it is a function on the configuration space so it depends only on the $PSL(2,\mathbb{R})$-class of ψ, $[\psi] \in \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$. A given spatial metric h_{ab} uniquely determines one such equivalence class $[\psi]$, as can be seen from App. B. There we explain an algorithm for constructing a ψ given a h_{ab}, but it contains an ambiguity associated with the conformal automorphisms of the unit complex disc which translates into a $PSL(2,\mathbb{R})$ indeterminacy, $\psi \sim \psi\chi$, where $\chi \in PSL(2,\mathbb{R})$. On the other direction, a class $[\psi]$ determines a spatial metric up to boundary-trivial conformal transformations, $[h] = [\Phi_* \Theta h]$, where $\Phi \in \mathit{Diff}^{+}(\Sigma)$ acts as the identity on the boundary and the function $\Theta \in C^\infty(\Sigma, \mathbb{R}^+)$ is 1 at the boundary. This can be seen from Sec. 6, where we explain how a given ψ determines the boundary value of the Weyl factor Λ from the condition on the boundary metric,

$$h_{ab}|_\partial = \psi_* \Lambda_\partial \bar{h}_{ab}|_\partial \tag{11.51}$$

so the only ambiguities are on extending $\psi \mapsto \Psi$ and $\Lambda_\partial \mapsto \Lambda$ to the interior of the disc, producing the spatial metric

$$h_{ab} = \Psi_* \Lambda \bar{h}_{ab} \tag{11.52}$$

that is, one could instead have chosen to extend these variables as $\psi \mapsto \Phi \circ \Psi$ and $\Lambda \mapsto \Theta\Lambda$, where (Φ, Θ) is boundary-trivial. Also, it is clear that could have started with $\psi\chi$ instead, and such a choice affects Λ_∂; however, it does it in such a way that the conformal transformation (Ψ, Ω) should be replaced by $(\Psi, \Omega) \circ (\Xi, \Gamma)$, where (Ξ, Γ) is a conformal isometry of $\bar{h}_{ab}$. Consequently, $(\Psi, \Omega) \circ (\Xi, \Gamma)$ defines the same metric h_{ab}. We thus conclude that the Q charges depend only on the conformal class of the spatial metric,

$$Q_n([\psi]) = Q_n([h_{ab}]) \tag{11.53}$$

with a slight abuse of notation.

It should be remarked that the fact that Q_n depends only on the (conformal class of the) spatial metric does not mean that Q_n cannot be expressed in a format that involves the extrinsic curvature. That is, K^{ab} could appear as long as it is combined with other quantities to produce a conformal invariant of h_{ab}. This is plausible only because the canonical charges are defined intrinsically on the reduced phase space, so any statements regarding their relationship with the ADM variables should be understood as being valid on-shell (i.e., assuming that the constraints hold), and consequently the extrinsic curvature is not independent from the spatial metric. In fact, it may well be that the extrinsic curvature appears in the most "natural form" of Q, i.e., the form with the most readily physical interpretation.

A second property, of Q_0 in particular, is that it must be bounded from above, attaining a maximum value of 2π when $[\psi] = [I]$, and unbounded from below [45]. In this configuration, $Q_n([I]) = \int d\theta \, e^{in\theta} = 2\pi\delta_{n0}$. This will be discussed in Sec. 12.4, where we shall see that the same property also holds in the quantum theory.

Lastly, we comment on an alternative point of view, based on the ADM analysis of corner symmetries, that may be helpful in clarifying the meaning of the Q charges. We shall refrain from going into much detail here since, to the extent that it applies to the present question, it is still an speculation.[56] The main point is to notice that, as proven in Sec. 4.2, the condition of fixing the induced boundary metric implies that all refoliations of the causal diamond can be achieved via *gauge* transformations. Moreover, we have also proved that, within the considered class of causal diamond spacetimes, a regular foliation always exists defined by the CMC condition, $K = -\tau$, which provides a universal gauge-fixing of time. We can think of this gauge-fixing

[56] Based on collaboration with Laurent Freidel, Luca Ciambelli and Ted Jacobson.

as introducing a constraint to the ADM phase space

$$C_\tau[\phi] := \int_\Sigma d^2x\, \phi(K+\tau) \tag{11.54}$$

where ϕ is a scalar labeling this family of functions. This can be seen as a constraint conjugate to the generator of refoliations, i.e., $C_\tau[\phi]$ and $H[N,\vec{0}]$, with $N|_\partial = 0$, form a family of second-class constraints. By themselves, generic corner-deforming ADM Hamiltonians $H[N,\vec{N}]$, where N and $\vec{N}$ are arbitrary at the boundary, would typically violate the boundary condition on the metric. But since $C_\tau[\phi]$ generates Weyl transformations on the spatial metric, we can always find linear combinations of $H[N,\vec{N}]$ and $C_\tau[\phi]$ that preserve the boundary condition. We conclude that the following Hamiltonian is symplectically differentiable and, for the appropriate choice of ϕ, generates a flow that preserves the boundary condition

$$H[N,\vec{N}] = H_{\text{bulk}}[N,\vec{N}] + C_\tau[\phi(N,\vec{N})] + 2\int_\partial ds\ \Big[N^c t_c t^a n^b K_{ab} - N^c n_c t^a t^b K_{ab} - NA\Big] \tag{11.55}$$

where A is the (scalar) acceleration of the corner as embedded in the spatial slice, i.e., $t^a \boldsymbol{\nabla}_a t^b = -An^b$.[57] There is another manner to write this expression which is interesting. From the perspective of the diamond spacetime, there are three directions that are special at the corner: one is along the corner itself, $(N,\vec{N}) = (0, \xi t^a)$,

[57]In a rigorous treatment of the second-class constraints in the ADM phase space, this Hamiltonian, with the appropriate constraint terms added, is precisely what should be used to compute Dirac brackets as regular Poisson brackets, i.e., $\{H_\partial, \cdot\}_D \approx \{H, \cdot\}$.

and the others are along the null rays of the diamond future horizon, $(N, \vec{N}) = (\zeta^+, -\zeta^- n^a)$, and past horizon, $(N, \vec{N}) = (-\zeta^-, -\zeta^- n^a)$. The charges associated with the horizon flows are related to the expansion parameters, $\Theta^{\pm}$, of the respective null generators; $\Theta^{\pm}$ is defined as the rate of change of length of a piece of corner, relative to its length, as it is transported along the null generators, $\pm u^a - n^a$. Therefore, we have three species of charges,

$$
\begin{aligned}
&\textit{Corner diffeomorphisms: } (N, \vec{N}) = (0, \xi t^a) \;\longrightarrow\; H[\xi] = 2 \int_\partial ds\, \xi t^a n^b K_{ab} \\
&\textit{Future horizon flow: } (N, \vec{N}) = (\zeta^+, -\zeta^+ n^a) \;\longrightarrow\; H[\zeta^+] = 2 \int_\partial ds\, \zeta^+ \Theta_+ \\
&\textit{Past horizon flow: } (N, \vec{N}) = (-\zeta^-, -\zeta^- n^a) \;\longrightarrow\; H[\zeta^-] = -2 \int_\partial ds\, \zeta^- \Theta_-
\end{aligned}
\tag{11.56}
$$

The charges associated with the corner diffeomorphisms are, as shown in 11.41, identified with (minus) P_ξ. We speculate that the Q charges are related to some (possibly non-linear) function of the expansion parameters, i.e., the charges associated with horizon flows. The investigation on this front is left to later work.

12 The quantum theory

The quantum theory is constructed based on a (projective) unitary irreducible representation of the transitive symmetry group $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$. Note that, in this case, quantization amounts to finding a class of suitable quantum theories, in which the (complete) subalgebra of canonical observers can be properly represented on the

Hilbert space — as explained in Sec. 9.1 this is unlike the case of a phase space with a vector space structure, where the quantization is based on the Heisenberg group which has a unique unitary irreducible representation (given a fixed value of $\hbar$).

Equivalently, as explained in App. F, projective unitary irreducible representations of a group are in one-to-one correspondence with (true) unitary irreducible representations of (the universal cover of) a central extension (by 2-cocycles) of the group, which in turn are in one-to-one correspondence with self-adjoint irreducible representations of the Lie algebra (of the extended group). Note however that, because of the Casimir matching principle discussed in Sec. 9.2, central elements of the algebra are to be represented with the same value as their classical counterparts, so if the Poisson algebra associated with the canonical group is truly homomorphic to the Lie algebra of the group, then we should not consider further central extensions of the algebra, even if in principle one is admissible. We showed in Sec. 11 that the Poisson algebra generated by the action of $\mathfrak{vira} \rtimes \mathit{Vira}$ reduces to

$$\begin{aligned}
\{P_n, P_m\} &= i(n-m)P_{n+m} \\
\{Q_n, P_m\} &= i(n-m)Q_{n+m} - 4\pi i n^3 \delta_{n+m,0} \\
\{Q_n, Q_m\} &= 0
\end{aligned} \tag{12.1}$$

so not only there is no need to extend $\mathfrak{vira} \rtimes \mathit{Vira}$ further, but the central charge associated with the *Vira* factor should be represented trivially. In this manner, even though we had to extend $\mathit{Diff}^+(S^1)$ into *Vira* for the purpose of constructing an appropriate transitive group of symplectomorphisms on the phase space, the algebra of configuration translations is reduced to the original $\mathfrak{diff}(S^1)$ algebra when realized as a Poisson algebra of momentum charges on the phase space, so the ultimate effect of that extension was only to introduce a central charge in the mixed bracket between momentum and configuration charges.

It is also worth noticing that the algebra above happens to be the $\mathfrak{bms}_3$ algebra of asymptotic symmetries at null infinity of asymptotically-flat spacetimes in three dimensions, as obtained in [43]. The group underlying this algebra is known as BMS_3, or the *Bondi–Metzner–Sachs group* in three dimensions. See that BMS_3 is not a subgroup of $\mathfrak{vira} \rtimes \mathit{Vira}$, but rather the latter is a central extension of the former.

In this way, the quantum theory is therefore based on some unitary irreducible representation of the "quantized" version of the classical Poisson algebra (12.1), in which the Poisson bracket $\{\,,\}$ is replaced by $-i[\,,]$,[58]

$$
\begin{aligned}
[\widehat{P}_n, \widehat{P}_m] &= (m-n)\,\widehat{P}_{n+m} \\
[\widehat{Q}_n, \widehat{P}_m] &= (m-n)\,\widehat{Q}_{n+m} + 4\pi n^3 \delta_{n+m,0} \\
[\widehat{Q}_n, \widehat{Q}_m] &= 0
\end{aligned}
\tag{12.2}
$$

where $\widehat{P}_n$ and $\widehat{Q}_n$ are operators on the Hilbert space $\mathcal{H}$ corresponding, respectively, to the observables P_n and Q_n. Note that they are not supposed to be self-adjoint since their classical counterparts are not real (as they are associated with complex Fourier modes of diffeomorphisms of S^1). Instead, these operators should satisfy the conjugation relations

$$
\begin{aligned}
(\widehat{P}_n)^\dagger &= \widehat{P}_{-n} \\
(\widehat{Q}_n)^\dagger &= \widehat{Q}_{-n}
\end{aligned}
\tag{12.3}
$$

which mimic the classical relations $(P_n)^* = P_{-n}$ and $(Q_n)^* = Q_{-n}$. We shall use the terminology that such a representation of the algebra is *unitary*, so that momentum and configuration variables labelled by real diffeomorphisms are self-

[58]Recall that we are using units where $\hbar = 1$. One would have to be careful in recovering $\hbar$ explicitly in these formulas. In particular, in Sec. 8.1, we have essentially borrowed the $\hbar$ from quantum mechanics in writing classical formulas for the metric h_{ab} and extrinsic curvature σ^{ab} in terms of the (dimensionless) quantities $\bar{h}_{ab}$ and $\bar{\sigma}^{ab}$. Had we not made the choice $\hbar = 1$, the physical symplectic form on the reduced phase space would actually be $\hbar\omega$, where ω is the (dimensionless) symplectic form associated with the cotangent bundle structure of $T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$. This $\hbar$ would appear in the definition of the canonical charges, and eventually cancel with the $1/i\hbar$ in the homomorphism to the quantum algebra.

adjoint operators and consequently the group of (real) symmetries is represented unitarily.

The unitary operators corresponding to the canonical transformations are obtained exponentiating the respective algebra elements. That is, we have

$$U(\exp(t(\widehat{\eta};0))) = e^{-it\widehat{Q}_{\widehat{\eta}}} \tag{12.4}$$

$$U(\exp(t(0;\widehat{\xi}))) = e^{-it\widehat{P}_{\widehat{\xi}}} \tag{12.5}$$

where t is a real parameter, exp denotes the Lie exponential in $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$ and elements of its Lie algebra are denoted by $(\widehat{\eta};\widehat{\xi}) \in \mathfrak{vira}^c \oplus \mathfrak{vira}$. The Casimir matching principle implies, from $T = 1$, that $\widehat{Q}_{\widehat{c}} = 1$, so

$$U(\exp(t(\widehat{c};0))) = U(((I,t);(I,0))) = e^{-it} \tag{12.6}$$

and also, from $R = 0$, that $\widehat{P}_{\widehat{c}} = 0$, so

$$U(\exp(t(0;\widehat{c}))) = U(((I,0);(I,t))) = 1 \tag{12.7}$$

where $I \in \mathit{Diff}^+(S^1)$. We have used here expression (10.6) parametrizing elements of *Vira* in terms of exponential of algebra elements.

If the representation is constructed directly at the group level, as in Sec. (12.1) below, then the configuration and momentum operators are obtained by differentiating the unitary operators,

$$\widehat{Q}_{\widehat{\eta}} = i \, \frac{\partial}{\partial t} U(\exp(t(\widehat{\eta}; 0)))\Big|_{t=0} \tag{12.8}$$

$$\widehat{P}_{\widehat{\xi}} = i \, \frac{\partial}{\partial t} U(\exp(t(0; \widehat{\xi})))\Big|_{t=0} \tag{12.9}$$

It automatically follows from the fact that U forms a representation of $\widetilde{G}$ that these P's and Q's satisfy the appropriate algebra, and from the unitarity that they will be self-adjoint (provided that the labels ξ and η are real diffeomorphisms).

12.1 Wavefunction realization

Let us now discuss the general representation theory for the canonical group $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$. Since this group has the form of a semi-direct product between a locally compact separable group $G = \mathit{Vira}$ and an abelian group $\mathfrak{vira}$, it is natural to employ Mackey's theory of induced representations [44]. (We offer a brief guide to Mackey's theory in App. E.) It should be stressed, nonetheless, that Mackey's theory formally only applies to finite dimensional groups, so we must be mindful that the existence, irreducibility and exhaustivity of the representations produced by Mackey's algorithm are not guaranteed (see however [46]). We note that the representation theory of BMS_3 has been studied recently [45, 104, 105, 106, 107], from Mackey's perspective, motivated by the important place of this group in the context of asymptotically-flat gravity in three dimensions.

According to Mackey's theory, an irreducible unitary representation of $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$ is characterized by a choice of an orbit $\mathcal{O}$ on $\mathfrak{vira}^*$ (generated by the dual action of Vira on $\mathfrak{vira}^*$, in this case the coadjoint action), together with a choice of an irreducible unitary representation of the little group $H_{\mathcal{O}}$ associated with $\mathcal{O}$ (i.e., $\mathcal{O} \sim \mathit{Vira}/H_{\mathcal{O}}$). The Hilbert space $\mathcal{H}_{\mathcal{O}}$ is realized as the space of sections Ψ of an associated bundle $\mathcal{S} \hookrightarrow \widetilde{G} \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$,[59] where $\mathscr{U} : H_{\mathcal{O}} \to \mathrm{Aut}(\mathcal{S})$ is an irreducible unitary representation of the corresponding little group $H_{\mathcal{O}}$ on a "little" Hilbert space $\mathcal{S}$. Informally, we can think of this associated bundle as a space defined by gluing a copy of $\mathcal{S}$ to each point of $\mathcal{O}$, and accordingly think of the quantum states Ψ as wavefunctions "living" on the orbit $\mathcal{O}$ valued in $\mathcal{S}$ — thus $\mathcal{S}$, if non-trivial, describes some sort of "internal states" or, in the language of particle physics, "intrinsic spin" degrees of freedom. We explain below how the group acts on this representation.

The group structure of $\mathfrak{vira} \rtimes \mathit{Vira}$, expressed in (10.57), is that Vira acts as the adjoint map on $\mathfrak{vira}$. In the notation of (E.12), $\delta : \mathit{Vira} \to \mathrm{Aut}(\mathfrak{vira})$ is

$$\delta_{\widehat{\psi}}\widehat{\eta} := \mathrm{ad}_{\widehat{\psi}}\widehat{\eta} \tag{12.10}$$

where $\widehat{\psi} \in \mathit{Vira}$ and $\widehat{\eta} \in \mathfrak{vira}$. The corresponding dual action on $\mathfrak{vira}^*$, denoted by

[59] A fiber bundle with total space E, base manifold M and fibers F is denoted by $F \hookrightarrow E \to B$. Note that the second arrow corresponds to the bundle projection map from E to B, and the first (hooked) arrow only indicates that F can be embedded into E as a fiber (although this embedding is not canonical).

$\widetilde{\delta}: \mathit{Vira} \to \mathfrak{vira}^*$, is precisely the coadjoint map,

$$\widetilde{\delta}_{\widehat{\psi}}\widetilde{\alpha} = \mathrm{coad}_{\widehat{\psi}}\widetilde{\alpha} \tag{12.11}$$

where $\widetilde{\alpha} \in \mathfrak{vira}^*$. Given an element $\widetilde{\alpha}_0 \in \mathfrak{vira}^*$, let $\mathcal{O}$ denote its orbit, i.e., the set of all points $\widetilde{\delta}_{\widehat{\psi}}\widetilde{\alpha}_0 = \mathrm{coad}_{\widehat{\psi}}\widetilde{\alpha}$ in $\mathfrak{vira}^*$, for all $\widehat{\psi} \in \mathit{Vira}$. The associated little group is

$$H_{\mathcal{O}} := \{\widehat{\chi} \in \mathit{Vira},\, \mathrm{coad}_{\widehat{\chi}}\widetilde{\alpha}_0 = \widetilde{\alpha}_0\} \tag{12.12}$$

Consequently, $\mathcal{O}$ is homeomorphic to $\mathit{Vira}/H_{\mathcal{O}}$. Now consider the principal bundle $H_{\mathcal{O}} \hookrightarrow \mathit{Vira} \to \mathcal{O}$, in which the projection map $q: \mathit{Vira} \to \mathcal{O} \sim \mathit{Vira}/H_{\mathcal{O}}$ is simply the group quotient, $q(\widehat{\psi}) = [\widehat{\psi}]$, where $[\widehat{\psi}] = [\widehat{\psi}\widehat{\chi}]$ for all $\widehat{\chi} \in H_{\mathcal{O}}$. More concretely, the quotient map can also be realized as $\widehat{\psi} \mapsto \mathrm{coad}_{\widehat{\psi}}\widetilde{\varepsilon}$. Given an irreducible unitary representation $\mathscr{U}: H_{\mathcal{O}} \to \mathrm{Aut}(\mathcal{S})$ of $H_{\mathcal{O}}$ on some vector space $\mathcal{S}$, the associated bundle $\mathcal{S} \hookrightarrow \mathit{Vira} \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$ is vector bundle over $\mathcal{O}$ with fibers $\mathcal{S}$, defined as follows: let the total space be the set of equivalence classes

$$\mathit{Vira} \times_{\mathscr{U}} \mathcal{S} := \Big\{[\widehat{\psi}, \varsigma] = [\widehat{\psi}\widehat{\chi}, \mathscr{U}(\widehat{\chi}^{-1})\varsigma];\ \text{where } \widehat{\psi} \in \mathit{Vira},\ \widehat{\chi} \in H_{\mathcal{O}} \text{ and } \varsigma \in \mathcal{S}\Big\} \tag{12.13}$$

and the projection map $q^{(\mathscr{U})}: \mathit{Vira} \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$ be

$$q^{(\mathscr{U})}([\widehat{\psi}, \varsigma]) := q(\widehat{\psi}) = [\widehat{\psi}] \tag{12.14}$$

This bundle has a natural linear structure defined by

$$[\widehat{\psi}, \lambda\varsigma + \lambda'\varsigma'] = \lambda[\widehat{\psi}, \varsigma] + \lambda'[\widehat{\psi}, \varsigma'] \tag{12.15}$$

Also, the *Vira* action on $\mathcal{O}$ has a natural lift to a *Vira* action on the associated bundle, denoted by $L : \mathit{Vira} \to \mathit{Diff}(\mathit{Vira} \times_{\mathscr{U}} \mathcal{S})$ and defined by

$$L_{\widehat{\phi}}[\widehat{\psi}, \varsigma] := [\widehat{\phi}\widehat{\psi}, \varsigma] \tag{12.16}$$

where $\widehat{\phi} \in \mathit{Vira}$. Finally, assume that $\mathcal{O}$ admits a measure μ that is quasi-invariant under *Vira*.[60] Representations defined for equivalent measures (i.e., having the same sets of measure zero) are unitarily equivalent, and under quite general conditions a homogeneous space admits a unique (up to equivalence) quasi-invariant measure (see footnote 92). In App. D we discuss a possible quasi-invariant measure for the orbit $\mathcal{O} = \mathcal{Q}$. The Hilbert space $\mathcal{H}$ is defined as the space of sections of the associated bundle,

$$\mathcal{H} := \Gamma(\mathit{Vira} \times_{\mathscr{U}} \mathcal{S}) := \{\Psi : \mathcal{O} \to \mathit{Vira} \times_{\mathscr{U}} \mathcal{S};\ \text{satisfying } q^{(\mathscr{U})}(\Psi(\widetilde{\alpha})) = \widetilde{\alpha} \text{ for all } \widetilde{\alpha} \in \mathcal{O}\} \tag{12.17}$$

[60]If $f : M \to N$ is a measurable map and μ is a measure on M, the push-forward of μ to N is defined by $f_*\mu[B] := \mu[f^{-1}(B)]$, where B is any Borel subset of N and f^{-1} designates the pre-image under f. A measure μ on a homogeneous space G/H is said to be *quasi-invariant* with respect to G if μ and its push-forwards $\mu_g := \delta_{g*}\mu$, for all $g \in G$, are equivalent to each other (i.e., have the same sets of measure zero).

with the inner product given by

$$\langle \Psi, \Psi' \rangle := \int_{\mathcal{O}} d\mu(\widetilde{\alpha}) \left\langle\!\left\langle \Psi(\widetilde{\alpha}), \Psi'(\widetilde{\alpha}) \right\rangle\!\right\rangle := \int_{\mathcal{O}} d\mu(\widetilde{\alpha}) \left(\varsigma(\widetilde{\alpha}), \varsigma'(\widetilde{\alpha}) \right) \tag{12.18}$$

where $\Psi(\widetilde{\alpha}) = [\widehat{\psi}(\widetilde{\alpha}), \varsigma(\widetilde{\alpha})]$, $\Psi'(\widetilde{\alpha}) = [\widehat{\psi}(\widetilde{\alpha}), \varsigma'(\widetilde{\alpha})]$ and $(\ ,\ )$ denotes the inner product in $\mathcal{S}$. (Note that Ψ and Ψ' must be expressed in terms of the same $\widehat{\psi} \in q^{-1}(\widetilde{\alpha})$ in this formula.) The irreducible unitary representation produced by this construction, $U : \mathfrak{vira} \rtimes \mathit{Vira} \to \mathrm{Aut}(\mathcal{H})$, is then

$$U(\widehat{\eta}, \widehat{\phi})\Psi(\widetilde{\alpha}) = e^{-i\widetilde{\alpha}(\widehat{\eta})} \sqrt{\frac{d\mu_{\widehat{\phi}}}{d\mu}(\widetilde{\alpha})}\, L_{\widehat{\phi}}\Big(\Psi(\mathrm{coad}_{\widehat{\phi}^{-1}} \widetilde{\alpha}) \Big) \tag{12.19}$$

where $(\widehat{\eta}, \widehat{\phi}) \in \mathfrak{vira} \rtimes \mathit{Vira}$ and $d\mu_{\widehat{\phi}}/d\mu$ is the Radon-Nikodym derivative associated with the measure μ. The form of this representation is compatible with our intuition in the sense that $\mathfrak{vira}$ corresponds classically to momentum translations and is quantum-mechanically represented as a pointwise phase rotation of the wavefunction, and *Vira* corresponds classically to configuration translations and is quantum-mechanically represented as translations of the wavefunction (i.e., notice that $L_{\widehat{\phi}}$ maps $\Psi(\mathrm{coad}_{\widehat{\phi}^{-1}} \widetilde{\alpha})$, which belongs to the fiber over $\mathrm{coad}_{\widehat{\phi}^{-1}} \widetilde{\alpha}$, to a point on the fiber over $\mathrm{coad}_{\widehat{\phi}}(\mathrm{coad}_{\widehat{\phi}^{-1}} \widetilde{\alpha}) = \widetilde{\alpha}$.)

Let us see how the conditions from the Casimir matching principle manifests here. Considering the group element $(t\widehat{c}; \widehat{I})$, where $t \in \mathbb{R}$ and $\widehat{I} = (I, 0)$ is the identity of *Vira*, formula (12.19) reads

$$U(t\widehat{c};\widehat{I})\Psi(\widetilde{\alpha}) = e^{-it\widetilde{\alpha}(\widehat{c})}L_{\widehat{I}}(\Psi(\widetilde{\alpha})) = e^{-it\widetilde{\alpha}(\widehat{c})}\Psi(\widetilde{\alpha}) \tag{12.20}$$

since $L_{\widehat{I}}$ acts trivially on the bundle. Comparing with (12.6) we conclude that

$$\widetilde{\alpha}(\widehat{c}) = 1 \tag{12.21}$$

that is, the orbit $\mathcal{O}$ must be chosen so that its central component is 1. Now consider the central element of *Vira*, $(0;(0,r))$, where $r \in \mathbb{R}$. Notice that since it acts trivially via the coadjoint map on $\mathfrak{vira}^*$, it always appear in $H_{\mathcal{O}}$ for any orbit. That is, $H_{\mathcal{O}}$ is always a central extension of the corresponding little group $K_{\mathcal{O}}$ of $\mathit{Diff}^+(S^1)$, $H_{\mathcal{O}} = K_{\mathcal{O}} \times_{ext} \mathbb{R}$, and moreover $\mathcal{O} \sim \mathit{Vira}/H_{\mathcal{O}} = \mathit{Diff}^+(S^1)/K_{\mathcal{O}}$. According to

(12.19) we have

$$\begin{aligned}
U(0;(0,r))\Psi(\widetilde{\alpha}) &= L_{(I,r)}(\Psi(\widetilde{\alpha})) \\
&= [(I,r)\widehat{\psi}(\widetilde{\alpha}), \varsigma(\widetilde{\alpha})] \\
&= [\widehat{\psi}(\widetilde{\alpha})(I,r), \varsigma(\widetilde{\alpha})] \\
&= [\widehat{\psi}(\widetilde{\alpha}), \mathscr{U}((I,r))\varsigma(\widetilde{\alpha})] \\
&= [\widehat{\psi}(\widetilde{\alpha}), e^{i\lambda r}\varsigma(\widetilde{\alpha})] \\
&= e^{i\lambda r}[\widehat{\psi}(\widetilde{\alpha}), \varsigma(\widetilde{\alpha})] \\
&= e^{i\lambda r}\Psi(\widetilde{\alpha}) \qquad (12.22)
\end{aligned}$$

where in the first line we used that $(0;(0,r))$ acts trivially on $\mathcal{O}$ (and, in particular, the Radon-Nikodym derivative is 1); in the second line we wrote $\Psi(\widetilde{\alpha}) = [\widehat{\psi}(\widetilde{\alpha}), \varsigma(\widetilde{\alpha})]$ (where, technically, $\widetilde{\alpha} \mapsto \widehat{\psi}(\widetilde{\alpha})$ is a *local* section) and used the definition (12.16) of the lifted action; in the third line we used that the central element commutes with any other; in the fourth line we used that $(0;(0,r)) \in H_{\mathcal{O}}$ so the equivalence relation (12.13) defining the associated bundle can be applied; in the fifth line we used that $(0;(0,r))$, being central, must be unitarily represented as $e^{i\lambda r}$ for some $\lambda \in \mathbb{R}$; in the sixth line we used the linear property (12.15) of the associated bundle; and in the last line we simply returned to the Ψ notation. The condition (12.7) thus implies that $\lambda = 0$,

$$\mathscr{U}((I,r)) = 1 \qquad (12.23)$$

i.e., the central factor of $H_{\mathcal{O}}$ must be represented trivially on $\mathcal{S}$. In summary, the Casimir matching principle applied to the central elements T and R implies that the representations of the canonical group should be restricted to those where the orbit $\mathcal{O}$ has unit central component in $\mathfrak{vira}^*$ and the central factor of $H_{\mathcal{O}} = K_{\mathcal{O}} \times_{ext} \mathbb{R}$ is trivially represented on $\mathcal{S}$. Formula (12.19) then reduces to

$$U((\eta, t); (\phi, r))\Psi(\widetilde{\alpha}) = e^{-it} e^{-i\alpha(\eta)} \sqrt{\frac{d\mu_\phi}{d\mu}(\widetilde{\alpha})}\, L_\phi\big(\Psi(\mathrm{coad}_{\phi^{-1}}\widetilde{\alpha})\big) \tag{12.24}$$

where $\widetilde{\alpha} = \alpha(\theta)d\theta^2 + \widetilde{c}$ and, as usual, $\alpha(\eta) = \int d\theta\, \alpha(\theta)\eta(\theta)$. Notice that now only $\phi \in \mathit{Diff}^+(S^1)$ (and not the central component of $\widehat{\phi}$) appears in the Radon-Nikodym derivative, L and coad.

From (12.8) and (12.9) we can compute the action of the quantized versions of the (non-central) canonical charges, $\widehat{Q}_\eta$ and $\widehat{P}_\xi$, on the wavefunctions. For the configuration charges we have

$$\begin{aligned} \widehat{Q}_\eta \Psi(\widetilde{\alpha}) &= i\frac{\partial}{\partial\lambda} U(\exp(\lambda(\eta, 0); 0)))\Psi(\widetilde{\alpha}) \\ &= i\frac{\partial}{\partial\lambda} e^{-i\lambda\alpha(\eta)}\Psi(\widetilde{\alpha}) \\ &= \alpha(\eta)\Psi(\widetilde{\alpha}) \end{aligned} \tag{12.25}$$

where the λ derivatives, here and next, are evaluated at $\lambda = 0$. For the momentum charges we have

$$\begin{aligned}\widehat{P}_\xi \Psi(\widetilde{\alpha}) &= i\frac{\partial}{\partial\lambda}U(\exp(\lambda(0;\xi)))\Psi(\widetilde{\alpha}) \\ &= i\frac{\partial}{\partial\lambda}\left[\sqrt{\frac{d\mu_{\exp(\lambda\xi)}}{d\mu}(\widetilde{\alpha})}L_{\exp(\lambda\xi)}\big(\Psi(\mathrm{coad}_{\exp(\lambda\xi)^{-1}}\widetilde{\alpha})\big)\right] \\ &= -i\mathscr{D}_\xi\Psi(\widetilde{\alpha}) \end{aligned} \tag{12.26}$$

where $\mathscr{D}_\xi$ is a derivative operator defined as

$$\mathscr{D}_\xi\Psi(\widetilde{\alpha}) := -\frac{\partial}{\partial\lambda}\left[\sqrt{\frac{d\mu_{\exp(\lambda\xi)}}{d\mu}(\widetilde{\alpha})}L_{\exp(\lambda\xi)}\big(\Psi(\mathrm{coad}_{\exp(\lambda\xi)^{-1}}\widetilde{\alpha})\big)\right]\Bigg|_{\lambda=0} \tag{12.27}$$

As expected from a derivative, this operator is linear under addition, $\mathscr{D}_\xi(\Psi+\Psi') = \mathscr{D}_\xi\Psi + \mathscr{D}_\xi\Psi'$, and satisfies a form of Leibniz rule under multiplication by scalars, $\mathscr{D}_\xi(f\Psi) = f\mathscr{D}_\xi + (X_\xi f)\Psi$, where $f : \mathcal{O} \to \mathbb{C}$ and X_ξ is the vector field on $\mathcal{O}$ induced by $\xi \in \mathfrak{diff}(S^1)$. Moreover, they form an anti-representation of $\mathit{Diff}^+(S^1)$, $[\mathscr{D}_\xi, \mathscr{D}_{\xi'}] = -\mathscr{D}_{[\xi,\xi']}$.

Finally, note that we should in principle consider the universal cover of the group in order to describe all projective representations. The group $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$ has fundamental group $\mathbb{Z}$, due to the *Vira* factor, which has been defined in Sec. 10.1 such that the fundamental group $\mathbb{Z}$ is inherited from $\mathit{Diff}^+(S^1)$. As described in App. D, the universal cover of $\mathit{Diff}^+(S^1)$, $\underline{\mathit{Diff}^+(S^1)}$, can be characterized as the space of diffeomorphisms of the real line, $\underline{\psi} : \mathbb{R} \to \mathbb{R}$, satisfying the condition

$$\underline{\psi}(\theta + 2\pi) = \underline{\psi}(\theta) + 2\pi \tag{12.28}$$

Thus, elements of the universal cover of *Vira*, $\underline{\mathit{Vira}}$, can be expressed as

$$\underline{\widehat{\psi}} = (\underline{\psi}, x) \in \underline{\mathit{Vira}} \tag{12.29}$$

The center $\mathcal{Z}$ of $\underline{\mathit{Diff}^+(S^1)}$ consists of functions $\underline{\psi}(\theta) = \theta + 2\pi n$, $n \in \mathbb{Z}$, all of which project to the identity in $\mathit{Diff}^+(S^1)$. Let us indicate the quotient $\underline{\mathit{Diff}^+(S^1)} \to \underline{\mathit{Diff}^+(S^1)}/\mathcal{Z} = \mathit{Diff}^+(S^1)$ simply by $\underline{\psi} \mapsto \psi$. Similarly, the center of $\underline{\mathit{Vira}}$ is $\mathcal{Z} \times \mathbb{R}$, i.e., consisting of elements $\underline{\widehat{\psi}} = (\theta + 2\pi n, r)$. We can describe the relationship between these two extensions (i.e., one by the 2-coclycle $\mathbb{R}$ and other by the fundamental group $\mathbb{Z}$) in terms of a commutative diagram

$$\begin{array}{ccc} \underline{\widehat{\psi}} \in \underline{\mathit{Vira}} & \xrightarrow{/\mathbb{R}} & \underline{\psi} \in \underline{\mathit{Diff}^+(S^1)} \\ {\scriptstyle /\mathbb{Z}}\downarrow & & \downarrow{\scriptstyle /\mathbb{Z}} \\ \widehat{\psi} \in \mathit{Vira} & \xrightarrow[/\mathbb{R}]{} & \psi \in \mathit{Diff}^+(S^1) \end{array}$$

where the arrows indicate group projections (or, in the reverse direction, group extensions). The diagram also helps elucidating the notation, i.e., see that "hat" denotes central extensions by 2-cocyles and "underline" denotes unwrapping by the fundamental group $\mathbb{Z}$ (the two accents are independent).

Evidently, the center $\mathcal{Z} \times \mathbb{R} \subset \underline{\mathit{Vira}}$ acts trivially through the coadjoint map from $\underline{Vira}$ to $\mathfrak{vira}^*$, so

$$\mathrm{coad}_{\underline{\widehat{\psi}}} = \mathrm{coad}_{\widehat{\psi}} = \mathrm{coad}_{\psi} \tag{12.30}$$

where the coadjoint map on the right-hand side is just the one from *Vira* to $\mathfrak{vira}^*$, given in (10.37). This means that the coadjoint orbits of $\underline{Vira}$ are exactly the same as the ones of *Vira*, the only difference being that the little group for $\underline{Vira}$ of an orbit $\mathcal{O}$, denoted by $\underline{H}_{\mathcal{O}}$, is some $\mathbb{Z}$-cover of the little group $H_{\mathcal{O}}$ for *Vira*. Therefore, the wavefunctions are still based on the same set of orbits, but they may carry "internal indices" in some *projective* unitary irreducible representations of the little group $H_{\mathcal{O}}$. In summary, the wavefunctions describing a projective unitary irreducible representations are sections of the bundle $\mathcal{S} \hookrightarrow \underline{Vira} \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$, where $\mathscr{U} : \underline{H}_{\mathcal{O}} \to \mathrm{Aut}(\mathcal{S})$ is a unitary irreducible representation of the little group $\underline{H}_{\mathcal{O}}$.

12.2 The monodromy and winding number Casimirs

If there are any Casimir invariants associated with the canonical algebra of observables, we can use the Casimir matching principle (explained at the end of Sec. 9.2) to filter out some representations. We have seen that the central T and R elements of $\widetilde{G}$, which are the simplest Casimir invariants, imply that the orbits must be restricted to those with central component equal to 1 ($\widetilde{\alpha} = \alpha(\theta)d\theta^2 + \widetilde{c}$) and the central $\mathbb{R}$ factor of the little group must be trivially represented. In this section we discuss a family of non-trivial Casimir invariants of $\mathfrak{bms}_3$ associated with the monodromy structure of coadjoint orbits of Virasoro.

Note that a classical observable that depends only on the Q_n charges, $C(\{Q_n\})$, can be unambiguously quantized to an operator $\widehat{C} := C(\{\widehat{Q}_n\})$, as there are no operator-ordering issues by virtue of the commutativity of the Q_n's. Furthermore, (12.25) implies that this operator acts on the wavefunctions by multiplication,

$$\widehat{C}\Psi(\widetilde{\alpha}) := C[\{\widehat{Q}_n\}]\Psi(\widetilde{\alpha}) = C[\{Q_n(\widetilde{\alpha})\}]\Psi(\widetilde{\alpha}) \tag{12.31}$$

where $Q_n(\widetilde{\alpha}) = \alpha(e^{in\theta}\partial_\theta) = \int d\theta\, e^{in\theta}\alpha(\theta) =: 2\pi\alpha_n$. These numbers α_n are of course just the Fourier components of $\widetilde{\alpha}$,

$$\widetilde{\alpha} = \sum_{n\in\mathbb{Z}} \alpha_n e^{-in\theta} d\theta^2 + \widetilde{c} \tag{12.32}$$

In particular, note that if we define the following $\mathfrak{diff}^*(S^1)$-valued operator,

$$\mathfrak{Q} := \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{Q}_n e^{-in\theta} d\theta^2 + \widetilde{c} \tag{12.33}$$

we get, in a formal sense,

$$\mathfrak{Q}\Psi(\widetilde{\alpha}) = \widetilde{\alpha}\Psi(\widetilde{\alpha}) \tag{12.34}$$

whose meaning is that, given any function $C : \mathfrak{diff}^*(S^1) \to \mathbb{R}$, we can associate to it a quantum operator $\widehat{C}$ defined by

$$\widehat{C} := C(\mathfrak{Q}) \tag{12.35}$$

and it follows that

$$\widehat{C}\Psi(\widetilde{\alpha}) := C(\widetilde{\alpha})\Psi(\widetilde{\alpha}) \tag{12.36}$$

Note that such an operator is defined in *any* irreducible representation of $\mathfrak{vira} \rtimes Vira$ in which $T = 1$.

If one finds a function $C : \mathfrak{diff}^*(S^1) \to \mathbb{R}$ that is constant on every orbit $\mathcal{O} \subset \mathfrak{vira}^*$, i.e.,

$$C(\widetilde{\alpha}) = C(\mathrm{coad}_\psi \widetilde{\alpha})\,, \ \text{ for all } \psi \in \mathit{Diff}^+(S^1) \tag{12.37}$$

which is therefore a classical Casimir observable (by restricting it to $\mathcal{O} = \mathcal{Q}$), then $\widehat{C} = C(\mathfrak{Q})$ will also be a quantum Casimir operator. In fact, a family of such functions exists [45] and is given by the trace of the k-th power of the *monodromy matrix*, $\mathbf{M}$. More explicitly, the function C_k is taken as

$$C_k(\widetilde{\alpha}) := Tr[\mathbf{M}^k] = Tr\left[\mathscr{P}\!\exp \int_0^{2\pi} d\theta\, k \begin{pmatrix} 0 & 1 \\ -\frac{\alpha(\theta)}{4} & 0 \end{pmatrix}\right] \tag{12.38}$$

where $k \in \mathbb{N}$, $\widetilde{\alpha} = \alpha(\theta)d\theta^2 + \widetilde{c}$, and Tr and $\mathscr{P}\!\exp$ are respectively the trace and the path ordered exponential (on the space of 2×2 matrices). We shall refer to this Casimir, which in the quantum theory is given by

$$\widehat{C}_k := C_k(\mathfrak{Q}) = Tr\left[\mathscr{P}\!\exp \int_0^{2\pi} d\theta\, k \begin{pmatrix} 0 & 1 \\ -\frac{\sum_{n\in\mathbb{Z}} \widehat{Q}_n e^{-in\theta}}{8\pi} & 0 \end{pmatrix}\right] \tag{12.39}$$

as the *Monodromy operator* of k-th order.[61]

Let us recall the how the monodromy matrix is defined. (See, e.g., [45] for details.) Associated to each element $\widetilde{\alpha} = \alpha(\theta)d\theta^2 + a\widetilde{c} \in \mathfrak{vira}^*$ there is a Hill's operator $\mathscr{H}_{\widetilde{\alpha}} := 4a\frac{\partial^2}{\partial\theta^2} + \alpha(\theta)$ acting on the space of real densities of weight $-1/2$ on $\mathbb{R}$ (i.e., objects that transform under diffeomorphisms like $F(\theta) = f(\theta)(d\theta)^{-\frac{1}{2}}$, where f is a scalar). The space of solutions of $\mathscr{H}_{\widetilde{\alpha}}F = 0$ forms a two-dimensional space, which can be displayed as a vector

$$\mathbf{F} := \begin{pmatrix} F_1 \\ \\ F_2 \end{pmatrix} \tag{12.40}$$

It is assumed that these solutions are normalized with respect to the Wronskian, $W(F_1, F_2) := F_1F_2' - F_2F_1' = -1$. These solutions are not 2π-periodic in general, but due to the 2π-periodicity of $\alpha(\theta)$ it follows that a 2π-translation corresponds to a change of basis in the space of solutions, i.e., there exists a matrix $\mathbf{M} \in \mathrm{SL}(2,\mathbb{R})$ such that

$$\mathbf{F}(\theta + 2\pi) = \mathbf{MF}(\theta) \tag{12.41}$$

This $\mathbf{M}$ is called the monodromy matrix associated with $\widetilde{\alpha}$ in the basis $\mathbf{F}$. A change of basis $\mathbf{F} \mapsto \mathbf{SF}$, where $\mathbf{S} \in \mathrm{SL}(2,\mathbb{R})$ (so that the Wronskian norm is preserved),

[61] To be more precise, $\widehat{C}$ is not a standard Casimir of $\mathfrak{vira} \rtimes \mathit{Vira}$, because T is replaced by the value 1. The "closest" one could get to defining a Casimir element of the universal enveloping algebra of $\mathfrak{vira} \rtimes \mathit{Vira}$ would be to replace $\frac{1}{8\pi}\sum_{n\in\mathbb{Z}}\widehat{Q}_n e^{-in\theta}$ by $\frac{1}{8\pi T}\sum_{n\in\mathbb{Z}}\widehat{Q}_n e^{-in\theta}$, which does not make sense since one cannot divide by the abstract symbol T. Nevertheless, this operator deserves to be called a "quantum Casimir" since it is the direct quantization of the classical Casimir $C[\{Q_n\}]$, is well-defined in any physically relevant irreducible representation (i.e., those where $T = 1$), and is represented as a multiple of the identity in each of these representations.

induces a change $\mathbf{M} \mapsto \mathbf{SMS}^{-1}$. Consequently, there is a basis-independent map from $\mathfrak{vira}^*$ to the space of conjugacy classes of $\mathrm{SL}(2,\mathbb{R})$, $\widetilde{\alpha} \mapsto [\mathbf{M}]$. Moreover, it is true that $\mathscr{H}_{\widetilde{\alpha}} F = 0 \Leftrightarrow \mathscr{H}_{\mathrm{coad}_\psi \widetilde{\alpha}}(\psi_* F) = 0$, for any $\psi \in \mathit{Diff}^+(S^1)$. It follows that this map from $\mathfrak{vira}^*$ to conjugacy classes of $\mathrm{SL}(2,\mathbb{R})$ is constant along each orbit $\mathcal{O}$ and, in particular, the traces of powers of $\mathbf{M}$ are real functions constant on the coadjoint orbits of Virasoro.

For an orbit with constant representative $\widetilde{\alpha}_0 = \alpha_0 d\theta^2 + \widetilde{c}$ this expression evaluates to

$$C_k(\widetilde{\alpha}_0) = 2\cos\left(k\pi\sqrt{\alpha_0}\right) \tag{12.42}$$

For the orbit $\mathcal{O} = \mathcal{Q}$, the representative is, as given in (10.43), $\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$, and therefore it is associated with the Casimir value

$$C_k(\widetilde{\varepsilon}) = 2(-1)^k \tag{12.43}$$

The Casimir matching principle thus restricts the possible quantum theories to those associated with orbits whose k-th power of the monodromy matrix has trace $2(-1)^k$. This leaves three possibilities for the $\mathrm{SL}(2,\mathbb{R})$ conjugacy classes of $\mathbf{M}$, whose representatives are

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}$$

It appears that there are no other *scalar* matrix-invariants that can be used to distinguish these cases.

It is possible to extend the Casimir matching principle to include other types of invariant objects, which we call *generalized Casimir operators*.[62] Let us define the *monodromy class operator*, $\mathfrak{M}$, valued in the space of conjugacy classes of $\mathrm{SL}(2,\mathbb{R})$, by

$$\mathfrak{M} := \left[\mathscr{P}\exp \int_0^{2\pi} d\theta \begin{pmatrix} 0 & 1 \\ -\frac{\sum_{n\in\mathbb{Z}} \widehat{Q}_n e^{-in\theta}}{8\pi} & 0 \end{pmatrix} \right] \tag{12.44}$$

This generalized operator is well-defined, both mathematically and physically: each entry of the 2×2 matrix is a function of the $\widehat{Q}_n$'s, which are commuting operators and can therefore be simultaneously diagonalized (or measured), and then the $\mathrm{SL}(2,\mathbb{R})$ conjugacy class of these "eigenmatrices" can be evaluated. In fact, the wavefunctions $\Psi(\widetilde{\alpha})$ are precisely the eigenvectors of this operator, and we can formally write

$$\mathfrak{M}\Psi(\widetilde{\alpha}) = [\mathbf{M}_{\widetilde{\alpha}}]\Psi(\widetilde{\alpha}) \tag{12.45}$$

[62] A possible formulation of a more generally applicable Casimir matching principle is as follows. Suppose that classically there is a function of the canonical observables, $C(\{H_{\widetilde{\xi}}\})$, valued in some space $\mathscr{S}$, which is constant on the phase space. Given any function $f : \mathscr{S} \to \mathbb{R}$, it is clear that $f \circ C$ is constant on the phase space and therefore it is classical Casimir observable. Say that C admits a natural quantization $\widehat{C}$ in the sense that, given any function $f : \mathscr{S} \to \mathbb{R}$, there is a naturally associated quantum operator $\widehat{C}_f := f \circ C(\widehat{H}_{\widetilde{\xi}})$ that commutes with all canonical observables, $[\widehat{C}_f, \widehat{H}_{\widetilde{\xi}}] = 0$, and is compatible with composition by real functions, $\widehat{C}_{\varphi\circ f'} = \varphi(\widehat{C}_{f'})$, where $\varphi : \mathbb{R} \to \mathbb{R}$. Each $\widehat{C}_f$ is therefore a quantum Casimir and must be irreducibly represented as a multiple of the identity. If, in a given representation, there exists $\mathfrak{c} \in \mathscr{S}$ such that $\widehat{C}_f = f(\mathfrak{c})$, for all f, then we say that $\widehat{C}$ has a $\mathscr{S}$-eigenvalue $\mathfrak{c}$ in that representation. It is then possible, in principle, to measure the $\mathscr{S}$-value of $\widehat{C}$, and the result of such measurement is the same for all quantum states, so we assume that this value should be matched with the classical $\mathscr{S}$-value of C.

The corresponding classical invariant is simply the $\mathrm{SL}(2,\mathbb{R})$ conjugacy class of the monodromy matrix, $\mathbf{M}$, which in the case of $\mathcal{Q} = \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$ evaluates to

$$[\mathbf{M}_{\mathcal{Q}}] = \left[\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right] \tag{12.46}$$

Thus we assume that the quantum theory should realize $\mathfrak{M} = [\mathbf{M}_{\mathcal{Q}}]$, which selects the orbits $\mathcal{O}$ whose monodromy class is

$$[\mathbf{M}_{\mathcal{O}}] = \left[\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right] \tag{12.47}$$

The only coadjoint orbits in this monodromy class are those with the following constant representatives [45]

$$\widetilde{\varepsilon}^{(n)} := n^2 d\theta^2 + \widetilde{c}\,, \text{ where } n \in 2\mathbb{Z}+1 \tag{12.48}$$

These orbits have topology $\mathit{Diff}^{+}(S^1)/\mathrm{PSL}^{(n)}(2,\mathbb{R})$, where $\mathrm{PSL}^{(n)}(2,\mathbb{R})$ is the subgroup of $\mathit{Diff}^{+}(S^1)$ generated by the subalgebra ∂_θ, $\sin(n\theta)\partial_\theta$ and $\cos(n\theta)\partial_\theta$.

There is another invariant associated with coadjoint orbits of Virasoro, the *winding number* [45]. This is a discrete parameter also associated with Hill's equation.

Given a (Wronskian-normalized) solution vector $\mathbf{F}$, define the map

$$\begin{aligned} &[0, 2\pi) \to \mathbb{RP}^1 \\ &\theta \mapsto [\mathbf{F}(\theta) \sim \lambda \mathbf{F}(\theta);\ \lambda \in \mathbb{R}] \end{aligned} \tag{12.49}$$

which should be seen as going from S^1 (the $[0, 2\pi)$ interval of $\mathbb{R}$) to S^1 (identified with $\mathbb{RP}^1$). The number of *complete* wrappings of the domain $[0, 2\pi)$ into the codomain $\mathbb{RP}^1$ induced by this map is called the winding number, w. By definition, $w \in \mathbb{Z}$. The winding number is independent of the basis of solutions $\mathbf{F}$ and invariant under the coadjoint action, $w_{\widetilde{\alpha}} = w_{\text{coad}_\psi \widetilde{\alpha}}$. Therefore, w is constant along each coadjoint orbit. Classically, the winding number of $\mathcal{Q}$ is 1. Quantum mechanically the winding number operator, $\mathfrak{w}$, can be constructed solely from the $\widehat{Q}_n$'s, and it acts on wavefunctions as

$$\mathfrak{w}\Psi(\widetilde{\alpha}) = w_{\widetilde{\alpha}}\Psi(\widetilde{\alpha}) \tag{12.50}$$

From the (generalized) Casimir principle, we conclude that the representations should be restricted to those with orbits with unit winding number,

$$w_{\mathcal{O}} = 1 \tag{12.51}$$

From the family of orbits in (12.48), the only candidate that has winding number 1 is $\widetilde{\varepsilon}^{(1)} = \widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$. That is, we conclude that the quantum theory should be based on the orbit $\mathcal{O} = \mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2, \mathbb{R})$, which we call the natural orbit.

12.3 The natural orbit and $PSL(2,\mathbb{R})$ indices

The natural choice for the orbit is

$$\mathcal{O} = \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R}) \tag{12.52}$$

so that the wavefunctions "live" on the configuration space (i.e., their domain is $\mathcal{Q}$). This was justified in the previous section from the Casimir matching principle, using the central element $T = 1$, the monodromy operator $\mathfrak{M}$ and winding number operator $\mathfrak{w}$. The little group is $K = PSL(2,\mathbb{R})$, so that the wavefunctions carry an internal index in some (projective) irreducible unitary representations of $PSL(2,\mathbb{R})$.

The representation theory of $PSL(2,\mathbb{R})$ is well-known [108, 109, 110, 111], so we briefly recapitulate it here to set up the notation. First note that since $PSL(2,\mathbb{R})$ is non-compact simple Lie group, the only finite-dimensional irreducible unitary representation is the trivial one [112]. The trivial representation would produce a quantization of the causal diamonds characterized by $\mathbb{C}$-valued wavefunctions living on $\mathcal{Q}$, which undoubtedly would be the "simplest guess" for the Hilbert space for the phase space $T^*\mathcal{Q}$. The other possibilities unveiled by the general quantization thus correspond to wavefunctions carrying infinitely many "internal states". The $\mathfrak{psl}(2,\mathbb{R})$ subalgebra is generated by the elements ∂_θ, $e^{i\theta}\partial_\theta$ and $e^{-i\theta}\partial_\theta$ of $\mathfrak{diff}(S^1)$, so in an irreducible unitary representation, $\mathscr{U}$, the respective self-adjoint generators,

υ_0, υ_+ and υ_-,[63] will satisfy the algebra

$$
\begin{aligned}
[\upsilon_0, \upsilon_+] &= \upsilon_+ \\
[\upsilon_0, \upsilon_-] &= -\upsilon_- \\
[\upsilon_-, \upsilon_+] &= 2\upsilon_0
\end{aligned}
\tag{12.53}
$$

There is one (independent) Casimir operator

$$
\zeta := \upsilon_0^2 - \frac{1}{2}\left(\upsilon_-\upsilon_+ + \upsilon_+\upsilon_-\right) \tag{12.54}
$$

which must be represented as a multiple of the identity,

$$
\zeta = \mu^2 - \frac{1}{4} \tag{12.55}
$$

where $\mu^2 \in \mathbb{R}$, and μ serves as a label for the representation. Now, as υ_0 is self-adjoint, we can assume that it has at least one eigenvector $|j;\mu\rangle$ (possibly in the weak sense, i.e., as a non-normalizable limit of a sequence of vectors) with eigenvalue $j \in \mathbb{R}$,

$$
\upsilon_0|j;\mu\rangle = j|j;\mu\rangle \tag{12.56}
$$

Since υ_0 generates the $SO(2)$ subgroup of $PSL(2,\mathbb{R})$, we must have, in a *true* rep-

[63]More precisely, $\mathscr{U}(\exp((c_0 + c_+e^{i\theta} + c_-e^{-i\theta})\partial_\theta)) = e^{-i(c_0\upsilon_0 + c_+\upsilon_+ + c_-\upsilon_-)}$.

resentation,

$$e^{i2\pi\upsilon_0} = 1 \tag{12.57}$$

so $j \in \mathbb{Z}$. We will later comment on its projective representations. From the algebra we see that $\upsilon_\pm$, if non-trivial, act as ladder operators for υ_0, i.e., $(\upsilon_\pm)^k|j;\mu\rangle \propto |j \pm k;\mu\rangle$. This is compatible with the previous observation that the spectrum of υ_0 is a subset of $\mathbb{Z}$. Using this ladder structure to build a basis we have

$$\upsilon_\pm|j;\mu\rangle = \left(j \pm \left(\mu + \frac{1}{2}\right)\right)|j \pm 1;\mu\rangle \tag{12.58}$$

where the vectors $|j;\mu\rangle$ are not necessarily normalized.[64] We have the following exhaustive list of possibilities:

1. *Trivial representation:* If $\mu = -1/2$ (so $\zeta = 0$) and $0 \in \mathit{Spectrum}(\upsilon_0)$, then all the generators annihilate $|0;\mu\rangle$. This is one-dimensional and, in fact, the only finite-dimensional unitary representation.

2. *Discrete series representation:* If $k := \mu + 1/2 \in \mathbb{Z} - \{0\}$, then notice that states with $j = \pm k$ will be annihilated by $\upsilon_\mp$. Thus, for each k, there are two distinct representations: one where $\upsilon_-|k;\mu\rangle = 0$, so $\mathit{Spectrum}(\upsilon_0) = \{k,\, k+1,\, \ldots\}$; and another where $\upsilon_+|-k;\mu\rangle = 0$, so $\mathit{Spectrum}(\upsilon_0) = \{\ldots,\, -k-1,\, -k\}$.

[64]Since the basis is discrete, these vectors can be normalized. By computing $\left\|\upsilon_+|j;\mu\rangle\right\|^2 = \langle j;\mu|\upsilon_-\upsilon_+|j;\mu\rangle$ with formula (12.58), one finds the relation

$$\left(j + \frac{1}{2} - \mu\right)\langle j;\mu|j;\mu\rangle = \left(j + \frac{1}{2} + \mu\right)\langle j+1;\mu|j+1;\mu\rangle$$

3. *Limit of discrete series representation:* If $k := \mu + 1/2 = 0$ and $0 \notin \mathit{Spectrum}(\upsilon_0)$, then there are two distinct representation: one where $\upsilon_-|1;\mu\rangle = 0$, so $\mathit{Spectrum}(\upsilon_0) = \{1,\, 2,\, \ldots\}$; and another where $\upsilon_+|-0;\mu\rangle = 0$, so $\mathit{Spectrum}(\upsilon_0) = \{\ldots,\, -2,\, -1\}$. Notice that, unlike the previous case, the spectrum does not start from k.

4. *Principal (spherical) series representation:* If $\mu^2 < 0$, then there is no minimal or maximum weight, i.e., $\mathit{Spectrum}(\upsilon_0) = \mathbb{Z}$. The representations for μ and $-\mu$ are equivalent.

5. *Complementary series representation:* If $\mu \in (-1/2, 0) \cup (0, 1/2)$ this also defines a representation in which $\mathit{Spectrum}(\upsilon_0) = \mathbb{Z}$. The reason that not all $\mu \in \mathbb{R} - \frac{1}{2}\mathbb{Z}$ is permissible is that, as can be seen from (12.58) and particularly footnote (64), some states $|j;\mu\rangle$ required to be included in the basis would have negative norm, so the representation would not be unitary. The representations for μ and $-\mu$ are equivalent.

If we consider projective representations of the canonical group, where the *Vira* factor is unwrapped to $\underline{\mathit{Vira}}$, the little group of the natural orbit $\mathcal{Q}$ would become $\underline{K} := \underline{PSL(2,\mathbb{R})}$, i.e., the universal cover of $PSL(2,\mathbb{R})$, where the $SO(2) \sim S^1$ subgroup of $PSL(2,\mathbb{R})$ is unwrapped to $\mathbb{R}$. In this way, the wavefunctions would also carry internal labels in *projective* representations of $PSL(2,\mathbb{R})$. The only modification in the analysis above is that condition (12.57) should not be imposed [111]. The consequence is that the spectrum of υ_0 is shifted by a fixed parameter, i.e., $\mathit{Spectrum}(\upsilon_0) = s + \mathbb{Z}$, where $s \in [0, 1)$. Thus, together with μ, s is another label classifying the unitary irreducible representations, which have the form

$$\begin{aligned} \upsilon_0|j;\mu,s\rangle &= (j+s)|j;\mu,s\rangle \\ \upsilon_\pm|j;\mu,s\rangle &= \left(j+s\pm\left(\mu+\frac{1}{2}\right)\right)|j\pm 1;\mu,s\rangle \end{aligned} \tag{12.59}$$

with j in (some subset of) $\mathbb{Z}$. The trivial representation obviously only exists when $s=0$, but the other unitary irreducible representations listed above are described in a completely analogous manner for each s, only with the appropriate shifts in the conditions for μ. For example, in the discrete series representation the condition would be $k=\mu+1/2\mp s\in\mathbb{Z}-\{0\}$, thus in one case $\upsilon_-|k;\mu,s\rangle=0$, so $\mathit{Spectrum}(\upsilon_0)=\{s+k,\, s+k+1,\ldots\}$; and in the other $\upsilon_+|-k;\mu,s\rangle=0$, so $\mathit{Spectrum}(\upsilon_0)=\{\ldots,\, s-k-1,\, s-k\}$. We anticipate that, from the analysis of the spin in Sec. 12.5, together with the assumption that the quantum theory incorporates a CMC time-reversal symmetry, we find that only the $s=0$ and $s=1/2$ cases are allowed. This corresponds to restricting only to the double cover of $PSL(2,\mathbb{R})$, i.e., $\mathrm{SL}(2,\mathbb{R})$.

Finally, let us discuss one simple manner to relate this representation of the little group with the canonical charges of the diamond, particularly, P_0, P_1 and P_{-1}. Given a wavefunction $\Psi(\widetilde{\alpha})$, let us study how the momentum operators act when evaluated at the conformal class of the identity, $[\psi]=[I]$, i.e., at the point $\widetilde{\varepsilon}=d\theta^2+\widetilde{c}$ of $\mathcal{Q}\subset\mathfrak{vira}^*$. Write $\Psi(\widetilde{\varepsilon})=[I,\varsigma]$, where $\varsigma\in\mathcal{S}$. The $PSL(2,\mathbb{R})$ subgroup of $\mathit{Diff}^{+}(S^1)$ stabilizes this point, so for any $\chi \in PSL(2,\mathbb{R})$, formula (12.24) gives

$$
\begin{aligned}
U(0;(\chi,0))\Psi(\widetilde{\varepsilon}) &= L_\chi(\Psi(\widetilde{\varepsilon})) \\
&= L_\chi[I,\varsigma] \\
&= [\chi,\varsigma] \\
&= [I,\mathscr{U}(\chi)\varsigma] \qquad (12.60)
\end{aligned}
$$

Writing $\chi = \exp[(c_0 + c_+e^{i\theta} + c_-e^{-i\theta})\partial_\theta]$ and differentiating with respect to the coefficients, we obtain

$$
\begin{aligned}
\widehat{P}_0\Psi(\widetilde{\varepsilon}) &= [I,\upsilon_0\varsigma] \\
\widehat{P}_1\Psi(\widetilde{\varepsilon}) &= [I,\upsilon_+\varsigma] \\
\widehat{P}_{-1}\Psi(\widetilde{\varepsilon}) &= [I,\upsilon_-\varsigma] \qquad (12.61)
\end{aligned}
$$

or, if we define $\Psi^{j;\mu,s}(\widetilde{\varepsilon}) := [I,|j;\mu,s\rangle]$, we have

$$
\begin{aligned}
\widehat{P}_0\Psi^{j;\mu,s}(\widetilde{\varepsilon}) &= (j+s)\Psi^{j;\mu,s}(\widetilde{\varepsilon}) \\
\widehat{P}_1\Psi^{j;\mu,s}(\widetilde{\varepsilon}) &= \left(j+s+\left(\mu+\frac{1}{2}\right)\right)\Psi^{j+1;\mu,s}(\widetilde{\varepsilon}) \\
\widehat{P}_{-1}\Psi^{j;\mu,s}(\widetilde{\varepsilon}) &= \left(j+s-\left(\mu+\frac{1}{2}\right)\right)\Psi^{j-1;\mu,s}(\widetilde{\varepsilon}) \qquad (12.62)
\end{aligned}
$$

and also

$$\widehat{\zeta}\Psi^{j;\mu,s}(\widetilde{\varepsilon}) = \left(\mu^2 - \frac{1}{4}\right)\Psi^{j;\mu,s}(\widetilde{\varepsilon}) \tag{12.63}$$

where

$$\widehat{\zeta} := \widehat{P}_0^2 - \frac{1}{2}\left(\widehat{P}_{-1}\widehat{P}_1 + \widehat{P}_1\widehat{P}_{-1}\right) \tag{12.64}$$

Note that $\widehat{\zeta}$ is not a Casimir for the canonical group $\widetilde{G}$, and therefore we cannot use the Casimir matching principle to deduce its value.

We emphasize that the relations above, connecting the action of the $\mathfrak{psl}(2,\mathbb{R})$ momenta to the little group representation $\mathscr{U}$, are valid on any wavefunction but only when evaluated at the specific point $\widetilde{\alpha} = \widetilde{\varepsilon}$ (i.e., $[\psi] = [I]$). Nevertheless, for wavefunctions localized at the conformal class of the identity, $[I]$, the formulas above are true operator actions (i.e., valid for all $\widetilde{\alpha} \in \mathcal{O}$). More precisely, consider the state $\Psi^{j;\mu,s}_{[I]}$ defined as

$$\Psi^{j;\mu,s}_{[I]}(\widetilde{\alpha}) := \begin{cases} \left[\widehat{\psi}(\widetilde{\alpha}), \delta\big(\widetilde{\alpha}, \widetilde{\varepsilon}\big)|j;\mu,s\rangle\right] & \text{for } \widetilde{\alpha} \in \mathcal{V} \\ 0 & \text{for } \widetilde{\alpha} \notin \mathcal{V} \end{cases} \tag{12.65}$$

where $\mathcal{V}$ is any open neighborhood of $[I] \in \mathcal{Q}$, $\widehat{\psi} : \mathcal{V} \to \mathit{Vira}$ is any local section of $PSL(2,\mathbb{R}) \times \mathbb{R} \hookrightarrow \mathit{Vira} \to \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ satisfying $\widehat{\psi}(\widetilde{\varepsilon}) = \widehat{I}$, and δ is a Dirac delta on $\mathcal{Q}$. This state is non-normalizable, so it must be understood as "generalized wavefunction".[65]

[65] See "Rigged Hilbert spaces" [113, 114].

12.4 The spectrum of $\widehat{Q}_0$

An interesting observable to discuss in the quantum theory is the Q_0 charge. Classically we have seen that it should correspond to some conformal invariant of the spatial metric (assuming the phase space constraints are satisfied). Moreover, it commutes with the spin P_0 and therefore is rotation-invariant in that sense. Quantum mechanically, in a representation based on the natural orbit, the spectrum of $\widehat{Q}_0$ is continuous, unbounded from below and bounded from above by 2π, as we explain in this section. Lastly, we remark that the spectrum of $\widehat{Q}_0$ will be used in the next section to evaluate the spectrum of the spin/twist, $\widehat{P}_0$.

In the context of asymptotically-flat spacetimes in 2+1 gravity, the group of asymptotic symmetries at the null infinity is BMS_3 [43]. The Q_0 charge would play the role of (minus) the total energy. The total energy generates an asymptotic diffeomorphism along the future null infinity, i.e., translations in the u parameter (where u is the retarded null coordinate, $u = t - r$).[66] It was shown by analysis of the wavefunction representation of $\mathfrak{bms}_3$ that the energy has a continuous spectrum, always unbounded from above and, in some representations, bounded from below. This is regarded as proof of energy positivity in 2+1 asymptotically-flat gravity [117].

[66] In the so-called BMS gauge, the (vacuum) metric takes the form $ds^2 = 8G\varepsilon(\theta)du^2 - 2dudr + 8G\left(j(\theta) + u\partial_\theta\varepsilon(\theta)\right)dud\theta + r^2d\theta^2$, where $\varepsilon(\theta)$ and $j(\theta)$ are interpreted as densities of energy and angular momentum, respectively. The total energy is $E = \frac{1}{2\pi}\int_0^{2\pi} d\theta\, \varepsilon(\theta)$. This reveals a quirk of asymptotically-flat gravity in 2+1 dimensions: the asymptotic diffeomorphism generated by the total energy E, ∂_u, is not necessarily timelike. In particular, for the Minkowski solution, ∂_u is timelike only if $E < 0$. When $E > 0$ the solutions correspond to the somewhat exotic "flat space cosmologies"[115, 116]. It is debatable whether E should still be called "energy" in the cases where it generates a spacelike flow.

We can straightforwardly import this result [45] to our theory of causal diamonds. Let us succinctly state the argument here. The action of $\widehat{Q}_0$ on a wavefunction $\Psi(\widetilde{\alpha})$ is expressed in (12.25) as

$$\widehat{Q}_0 \Psi(\widetilde{\alpha}) = \alpha(\partial_\theta)\Psi(\widetilde{\alpha}) \tag{12.66}$$

and, in the characterization $\widetilde{\alpha} = \mathrm{coad}_\psi \widetilde{\varepsilon} = [\psi]$, it follows from (10.45) that

$$\widehat{Q}_0 \Psi([\psi]) = \int_0^{2\pi} d\theta \, \frac{1 - 2S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} \Psi([\psi]) \tag{12.67}$$

Therefore, any properties that are classically satisfied by Q_0 on the orbit $\mathcal{O} = \mathcal{Q}$, will also be satisfied quantum-mechanically. (In fact, this is true for any function of the Q charges since they act as multiplication on the wavefunctions.) The property of interest here is a general inequality involving the Schwarzian, known as the average lemma [98, 74, 118, 45], which reads

$$\int_0^{2\pi} d\theta \, S[\psi](\theta) \leq \int_0^{2\pi} d\theta \, \frac{1 - \psi'(\theta)^2}{2} \tag{12.68}$$

which is saturated if and only if $\psi \in PSL(2,\mathbb{R})$. Now let us replace ψ be ψ^{-1} in this inequality, and rewrite it as

$$\int_0^{2\pi} d\theta \, \left({\psi^{-1}}'(\theta)^2 + 2S[\psi^{-1}](\theta) \right) \leq \int_0^{2\pi} d\theta = 2\pi \tag{12.69}$$

Using relation (10.24) to express $S[\psi^{-1}]$ in terms of $S[\psi]$, and also ${\psi^{-1}}'(\theta) =$

$1/\psi'(\psi^{-1}(\theta))$, we get

$$\int_0^{2\pi} d\theta\, \frac{1 - 2S[\psi](\psi^{-1}(\theta))}{\psi'(\psi^{-1}(\theta))^2} \leq 2\pi \tag{12.70}$$

From the formula for Q_0 above, it follows that the possible eigenvalues are less or equal to 2π. Moreover, it is clear that any real value below 2π is attainable by at least one ψ.[67] Therefore, we have

$$Spectrum(\widehat{Q}_0) = (-\infty, 2\pi] \tag{12.71}$$

While most eigenspaces of $\widehat{Q}_0$ are highly degenerate, as there are many diffeomorphism classes $[\psi]$ on which Q_0 takes the same value, there is one eigenspace that is special: the one associated to the maximum eigenvalue 2π. From the average lemma, the inequality is saturated only when $\psi \in PSL(2,\mathbb{R})$, that is, when $[\psi] = [I]$. This implies that the (non-normalizable) wavefunctions $\Psi^{j;\mu}_{[I]}$, introduced in 12.65, are the only (generalized) states in which $\widehat{Q}_0 = 2\pi$,

$$\widehat{Q}_0 \Psi^{j;\mu}_{[I]} = 2\pi \Psi^{j;\mu}_{[I]} \tag{12.72}$$

Moreover, by direct evaluation, we see that these states are annihilated by all other $\widehat{Q}_n$'s,

$$\widehat{Q}_n \Psi^{j;\mu}_{[I]} = 0\,, \text{ for } n \neq 0 \tag{12.73}$$

[67] For example, consider $\psi(\theta) = \theta + \frac{\kappa}{n}\sin(n\theta)$, where $n \in \mathbb{Z}$ and $0 < \kappa < 1$. This describes a valid diffeomorphism of S^1 since $\psi'(\theta) = 1 + \kappa\cos(n\theta) > 0$. If $\kappa \ll 1$, then $Q_0([\psi^{-1}]) = \int_0^{2\pi} d\theta \left(2S[\psi](\theta) + \psi'(\theta)^2\right) \approx -3\pi(\kappa n)^2 + 2\pi$. Since $(\kappa n)^2$ can be any positive real number, $Q_0([\psi^{-1}])$ can be arbitrarily negative.

Classically, these states correspond a causal diamonds whose spatial geometries are in the same conformal class as the symmetric disc.

12.5 Spin/Twist quantization

The quantum theory reveals a quite interesting result: the spin, which we have interpreted as the twist of the boundary loop, is quantized. To see this, note that $\widehat{P}_n$ and $\widehat{Q}_n$ satisfy the following commutation relations with $\widehat{P}_0$,

$$[\widehat{P}_0, \widehat{P}_n] = n\widehat{P}_n$$
$$[\widehat{P}_0, \widehat{Q}_n] = n\widehat{Q}_n \tag{12.74}$$

We are interested in representations in which $\widehat{P}_0$ is self-adjoint (since $\widehat{P}_0^\dagger = \widehat{P}_{-0}$), so it must have a real spectrum. Let us say that $\widehat{P}_0$ has an eigenvector $|s\rangle$ with eigenvalue $s \in \mathbb{R}$; even though, rigorously, it is not guaranteed a priori that $\widehat{P}_0$ has any eigenvectors/eigenvalues. The commutation relations above imply that both $\widehat{P}_n$ and $\widehat{Q}_n$ act as ladder operators for $\widehat{P}_0$, raising its eigenvalues by n. That is, $\widehat{P}_n|s\rangle$ and $\widehat{Q}_n|s\rangle$, if non-zero, are eigenvectors of $\widehat{P}_0$ with eigenvalue $s+n$. Since the canonical algebra must be represented irreducibly, the whole Hilbert space is spanned by acting with all $\widehat{P}_n$'s and $\widehat{Q}_n$'s (and their products) on the any given state, such as $|s\rangle$. Also, note that any string of P's and Q's also acts as a ladder operator for $\widehat{P}_0$, raising the eigenvalues by some integer: for example, $\widehat{P}_1(\widehat{P}_3)^2\widehat{Q}_{-4}\widehat{P}_2$ raises spin eigenvalues by $1 + 2 \times 3 - 4 + 2 = 5$. Therefore we conclude that the spectrum of $\widehat{P}_0$ is some subset of $\{s+n; n \in \mathbb{Z}\}$, where s is some (fixed) real number.[68] Next we show that the spectrum of $\widehat{P}_0$ is in fact equal to $s+\mathbb{Z}$, that is, there are no level gaps.

First, note that since the spectrum of $\widehat{P}_0$ is discrete (i.e., consisting of isolated points), it must be that $\widehat{P}_0$ is diagonalizable by (normalizable) eigenvectors, whose eigenvalues coincide with the spectrum [84, 119, 120]. Let $|\Psi_j\rangle$ be a (normalizable) eigenvector of $\widehat{P}_0$ with eigenvalue $s+j$, for some $j \in \mathbb{Z}$. Now *suppose* that either of the values $s+j+1$ or $s+j-1$ do not belong to the spectrum of $\widehat{P}_0$, and let us show that this leads to a contradiction. We consider the case where $s+j-1$ is assumed not to be in the spectrum, as the other case is analogous. We should have

$$\widehat{Q}_{-1}|\Psi_j\rangle = \widehat{P}_{-1}|\Psi_j\rangle = 0 \tag{12.75}$$

From the algebra of $\widehat{P}$ and $\widehat{Q}$, as in (12.2), we have

$$[\widehat{Q}_1, \widehat{P}_{-1}] = 2\big(2\pi - \widehat{Q}_0\big) \tag{12.76}$$

and contracting with $|\Psi_j\rangle$ we get

$$\langle \widehat{Q}_{-1}\Psi_j|\widehat{P}_{-1}\Psi_j\rangle - \langle \widehat{P}_1\Psi_j|\widehat{Q}_1\Psi_j\rangle = 2\langle\Psi_j|2\pi - \widehat{Q}_0|\Psi_j\rangle \tag{12.77}$$

[68] At this stage, s should not be confused with the shift parameter of the little group representation described in Sec. 12.3. But, as the notation suggests, they will be eventually related.

where we used that $\langle\Psi_j|\widehat{Q}_1 = \langle\Psi_j|\widehat{Q}^\dagger_{-1} = \langle\widehat{Q}_{-1}\Psi_j|$. In Sec. 12.4 we showed that the spectrum of $\widehat{Q}_0$, for the natural orbit $\mathcal{O} = \mathcal{Q}$, is bounded from above by 2π, only attaining the maximum on the non-normalizable states concentrated at the conformal class of the identity, $\Psi_{[I]}$. Therefore,

$$\langle\Psi_j|2\pi - \widehat{Q}_0|\Psi_j\rangle > 0 \tag{12.78}$$

since Ψ_j is normalizable (and thus not equal to $\Psi_{[I]}$). Most importantly, this expectation value is non-zero. Hence, if (12.75) is true,

$$\langle\widehat{P}_1\Psi_j|\widehat{Q}_1\Psi_j\rangle = -2\langle\Psi_j|2\pi - \widehat{Q}_0|\Psi_j\rangle \neq 0 \tag{12.79}$$

implying, in particular, that

$$\widehat{Q}_1|\Psi_j\rangle \neq 0 \tag{12.80}$$

On the other hand, note that

$$\begin{aligned}\langle\widehat{Q}_1\Psi_j|\widehat{Q}_1\Psi_j\rangle &= \langle\Psi_j|\widehat{Q}_{-1}\widehat{Q}_1|\Psi_j\rangle \\ &= \langle\Psi_j|\widehat{Q}_1\widehat{Q}_{-1}|\Psi_j\rangle \\ &= 0\end{aligned} \tag{12.81}$$

where we used the commutativity of the $\widehat{Q}$'s. From the non-degeneracy of the Hilbert space metric it follows that

$$\widehat{Q}_1|\Psi_j\rangle = 0 \tag{12.82}$$

which is in contradiction with (12.80). Therefore we conclude that (12.75) is false, so the value $j+s-1$ must belong to the spectrum. The same argument applies to any value $j+s$ in the spectrum, implying that its two neighbors are also in the spectrum. We have thus proved that the spectrum of $\widehat{P}_0$, in any non-trivial unitary irreducible representation based on $\mathcal{O} = \mathcal{Q}$, is

$$\mathit{Spectrum}(\widehat{P}_0) = s + \mathbb{Z} \tag{12.83}$$

Without loss of generality, s can be taken to be in the interval $[0,1)$.

The value of s can be restricted by assuming time-reversal symmetry. That is, if one wishes to implement the CMC time-evolution Hamiltonian (discussed in Sec. 7.2) into the quantum theory, in such a way that its classical time-reversal symmetry $\tau \mapsto -\tau$ is preserved,[69] then there must exist an anti-unitary operator $\mathfrak{T}$

[69]Under time-reversal the ADM variables change as $h_{ab} \mapsto h_{ab}$, $\sigma^{ab} \mapsto -\sigma^{ab}$ and $\tau \mapsto -\tau$. This is because the induced metric does not depend on the time orientation but the extrinsic curvature $K_{ab} = \boldsymbol{\nabla}_a u_b$ is defined with u being "future directed", which flips sign when time is reversed. In terms of the conformal variables this translates into $\psi \mapsto \psi$, $\bar{\sigma}^{ab} \mapsto -\bar{\sigma}^{ab}$ and $\tau \mapsto -\tau$. The time-evolution Hamiltonian, related to the area of the CMC slice, depends only on the solution λ to the Lichnerowicz equation which is not sensitive to this change of signs. This transformation is an anti-symplectic symmetry. Finally note that, classically, P_n depends linearly on σ^{ab}, so it changes sign, and Q_n depends only on ψ which does not change.

acting on the canonical algebra as

$$\mathfrak{T}\widehat{P}_n\mathfrak{T}^\dagger = -\widehat{P}_n$$
$$\mathfrak{T}\widehat{Q}_n\mathfrak{T}^\dagger = \widehat{Q}_n \tag{12.84}$$

Naturally, this implies that $\mathfrak{T}^\dagger|s\rangle$ is an eigenvector of $\widehat{P}_0$ with eigenvalue $-s$, so the spectrum of $\widehat{P}_0$ must be symmetric under a change of sign. Therefore, s must be integer or half-integer, so the spectrum of $\widehat{P}_0$ must be either $\mathbb{Z}$ or $\frac{1}{2}+\mathbb{Z}$.

The value of s is naturally related to the representation of the little group in the wavefunction realization. In the representation based on the natural orbit, $\mathcal{O} = \mathcal{Q}$, the wavefunctions carry internal indices associated with (projective) unitary irreducible representations of $PSL(2,\mathbb{R})$, as described in 12.3. In particular, consider the state $\Psi_{[I]}$ defined in (12.65). This state is a (generalized) eigenvector of $\widehat{P}_0$, for according to (12.62) we have

$$\widehat{P}_0\Psi^{j;\mu}_{[I]} = (j+s)\Psi^{j;\mu}_{[I]} \tag{12.85}$$

where j some integer. Therefore, we see that the shift parameter characterizing the spectrum of $\widehat{P}_0$ is the same s parameter labeling the (projective) representation of the little group $PSL(2,\mathbb{R})$. In particular, notice that the restriction to $s=0$ or $s=1/2$, imposed by the time-reversal symmetry, implies that we need to consider only true representations of the double cover of $PSL(2,\mathbb{R})$, that is, $\mathrm{SL}(2,\mathbb{R})$. The case $s = 0$ is possibly the most natural, associated with a "bosonic" diamond, while $s = 1/2$ corresponds to a "fermionic" diamond (in the sense that a rotation by 2π flips the sign of the wavefunction). Lastly, note that all values of $j \in \mathbb{Z}$ are contained in the spectrum of $\widehat{P}_0$, as proven directly from the canonical algebra, even if the representation of $PSL(2,\mathbb{R})$ is one of those that only include a subset of integers (this can be interpreted as the fact that P_0 is the "total spin", which is not entirely accounted for by the "internal spin" but also receives a contribution from the "orbital angular momentum" of the wavefunction).

Combining the quantization condition above with formula (11.50), relating the spin P_0 with the twist $\mathcal{T}$ of the boundary loop, yields a quantization condition for the twist

$$\mathcal{T} = \frac{16\pi^2 \ell_P}{\ell}(n+s)\,, \quad n \in \mathbb{Z} \tag{12.86}$$

where $s = 0$ or $1/2$. Therefore we see that the twist can only change in discrete increments defined by the ratio of the Planck length to the boundary length. In particular, this is consistent with our intuition about the "classical limit" in the sense that the size of this increment goes to zero as $\ell \gg \ell_P$, so a classical diamond would be in a state with a very large number of "twist quanta".

Now let us determine if our choice of boundary metric $\gamma = (\ell/2\pi)^2 d\theta^2$ was merely a convention, or whether it has true physical implications. Note that $P_0 := P_{\partial_\theta}$ is defined with respect to the reference unit disc, so the choice of $\gamma = \gamma(\theta)d\theta^2$ will affect its geometrical interpretation. In particular, the tangent vector to the boundary, t^a, with unit norm with respect to the physical metric γ, would be generally given by

$$t^a|_\theta = \frac{1}{\sqrt{\gamma(\theta)}}\partial_\theta \tag{12.87}$$

which of course reduces to (11.32) when $\gamma(\theta) = (\ell/2\pi)^2$. Accordingly, formula (11.33) is updated to

$$P_\xi = -\frac{\ell}{16\pi^2 \ell_P}\int_\partial ds\, K_{ab} n^a \frac{2\pi\sqrt{\gamma}\xi}{\ell} t^b \tag{12.88}$$

where $\xi = \xi(\theta)\partial_\theta$. Thus, the charge generating isometric rotations of the boundary, the physical spin S, corresponding to a $\xi^a \propto t^a$ with a constant coefficient as given in (11.42), will not in general be equal to P_0, but instead

$$S := -\frac{\ell}{16\pi^2 \ell_P}\int_\partial ds\, K_{ab} n^a t^b = P_{\xi_\gamma} \tag{12.89}$$

where

$$\xi_\gamma := \frac{\ell}{2\pi\sqrt{\gamma}}\partial_\theta \tag{12.90}$$

In particular, $S = P_0$ only when $\gamma(\theta) = (\ell/2\pi)^2$. The defining characteristic of ξ_γ is that it is a nowhere vanishing vector field on the boundary with total parameter-length ℓ. In fact, for any two choices of boundary metric, γ and γ', with the same total length ℓ, there exists a diffeomorphism $\psi : S^1 \to S^1$ such that

$$\gamma' = \psi^* \gamma \tag{12.91}$$

and, consequently,

$$\xi_{\gamma'} = \psi_* \xi_\gamma \tag{12.92}$$

In particular, if with one choice of boundary metric the spin gets associated with the operator $\widehat{P}_\xi$, then with another choice it would be associated with the operator $\widehat{P}_{\psi_*\xi}$ instead. The specific choice of the boundary metric, beyond the information contained in the total length ℓ, would therefore be physically meaningful if the spectrum of $\widehat{P}_{\psi_*\xi}$ were different than that of $\widehat{P}_\xi$. We now show that this is not the case. That is, we will show that there is a unitary map relating $\widehat{P}_{\psi_*\xi}$ with $\widehat{P}_\xi$, so they must have the same spectrum. In fact, as one could naturally guess, the unitary map generating this transformation is associated with the canonical group action by the element $(0; \widehat{\psi}) \in (\mathfrak{vira}^*)^* \rtimes Vira$, where $\widehat{\psi} = (\psi, r)$ (in fact the central component of $\widehat{\psi}$ does not matter since it is represented trivially). Take formula (12.5) defining $\widehat{P}_\xi$ and sandwich it between $U(0; \widehat{\psi})$ and its adjoint,

$$U((0; \widehat{\psi}))U(\exp(t(0; \xi)))U((0; \widehat{\psi}))^\dagger = U((0; \widehat{\psi}))e^{-itP_\xi}U((0; \widehat{\psi}))^\dagger \tag{12.93}$$

On the left-hand side we have

$$\begin{aligned}
U((0;\widehat{\psi}))U(\exp(t(0;\xi)))U((0;\widehat{\psi}))^\dagger &= U((0;\widehat{\psi})\exp(t(0;\xi))(0;\widehat{\psi})^{-1}) \\
&= U(\exp(t\,\mathrm{ad}_{(0;\widehat{\psi})}(0;\xi))) \\
&= U(\exp(t(0;\psi_*\xi + \Lambda_\psi(\xi)\widehat{c}))) \\
&= U(\exp(t(0;\psi_*\xi)))U(\exp(t(0;\Lambda_\psi(\xi)\widehat{c}))) \\
&= U(\exp(t(0;\psi_*\xi))) \\
&= e^{-it\widehat{P}_{\psi_*\xi}} \qquad (12.94)
\end{aligned}$$

where on the first line we used the definition of a representation; on the second line that, for any Lie group, $\mathrm{Ad}_g \exp(\xi) = \exp(\mathrm{ad}_g \xi)$; on the third line expression (10.18) for the adjoint action on *Vira*; on the fifth line we factored out the central piece; on the sixth line we used that the central element of *Vira* is represented trivially; and on the last line we used again (12.5). On the right-hand side we have

$$U((0;\widehat{\psi}))e^{-it\widehat{P}_\xi}U((0;\widehat{\psi}))^\dagger = e^{-itU((0;\widehat{\psi}))\widehat{P}_\xi U((0;\widehat{\psi}))^\dagger} \qquad (12.95)$$

so we conclude

$$U((0;\widehat{\psi}))\widehat{P}_\xi U((0;\widehat{\psi}))^\dagger = \widehat{P}_{\psi_*\xi} \qquad (12.96)$$

as we intended to show.

13 Conclusion

In this work we have explored the canonical quantization of causal diamonds in (2+1)-dimensional gravity, with a non-positive cosmological constant, via the reduced phase space approach, in conjunction with Isham's group-theoretic quantization method. In this discussion section we recapitulate on what has been achieved, from a general perspective, and comment on possible future directions to explore.

As explained in the introduction (Sec. 1), a motivation for this project was to understand how causal diamonds, which serve as the prototype notion of subsystems in classical gravity, could be described in quantum mechanics. The concept of a subsystem lies at the core of the current philosophical framework of physics, in which one imagines the universe as consisting of a net of subsystems, such that in some regimes certain "pieces of the universe" can be satisfactorily described independently from the "rest of the universe", and if two such subsystems are allowed to interact then their union becomes the new subsystem under consideration. Naturally this philosophy is intimately related to the principle of locality and causality, where causally-disconnected regions of spacetime can be treated as independent subsystems. Nevertheless, quantum gravity points to the necessity of a fundamental revision of this concept due to the absence of compactly-supported gauge-invariant observables. There are two plausible directions to address this point: one would be to quantize gravity in the whole universe and then try to figure out what is the net of subsystems at the quantum level (e.g., by trying to find some "order" in the highly intricate algebraic structure of observables); and the other would be to quantize the classical gravitational notion of subsystems (i.e., causal diamonds) and use the resulting quantum objects as building blocks to construct the theory of quantum gravity in the whole universe. It is unclear which approach is more promising, and whether they would be equivalent in any sense, but here we chose to investigate the latter. What we have achieved is only a first step in this program, as we have described a fully non-perturbative kinematical quantization of causal diamonds by treating them as a self-contained system. A next step would be to consider how (or if) quantum causal diamonds can be sewn together to assemble the entire spacetime.

Aligned with the goal of implementing a rigorous, non-perturbative quantization of a gravitational system, a second motivation we had was to continue the exploration of Moncrief's program of quantizing gravity. In this program, one employs a convenient gauge-fixing of time where the spacetime is foliated by surfaces of constant-mean-curvature. This provides a surprisingly general prescription to solve the constraints of gravity in a variety of cases (i.e., those admitting such a foliation), removing the associated gauge ambiguities, and ending with a reduced phase space equal to the cotangent bundle of an appropriate space of conformal geometries on the Cauchy slice. Most applications of this idea have been to systems with a closed Cauchy slice, and here we applied it to causal diamonds whose Cauchy slice has a finite boundary. We made some simplifying assumptions, so that the problem could be treated more concretely, such as considering pure gravity (i.e., General Relativity with no matter fields), lower spacetimes dimension (i.e., 2+1) and a trivial topology for the Cauchy slice (i.e., a disc). It could be interesting to extend the formalism to more general causal diamonds, in higher dimensions, with matter and non-trivial topologies. We have also considered that the (induced) metric on the boundary is fixed. The only intrinsic parameter characterizing this condition is the total length ℓ of the corner, in the sense that two choices of boundary metrics with the same total length are related by a symplectomorphism at the classical level and, according to our quantization, a unitary transformation at the quantum level. This condition was introduced so that the CMC gauge would be attainable (i.e., so that an arbitrary Cauchy slice could always be deformed into a CMC surface), and for that reason it is not clear whether this boundary condition could be altered or removed. Lastly, we have decided to carry out the reduction of the phase space from the perspective of the ADM formulation, described in terms of spatial metrics and extrinsic curvatures, in part because the boundary condition on the metric was expressed very naturally. It could also be interesting to investigate the phase space reduction from the perspective of the Chern-Simons formulation, which is described in terms of a pair of $\mathrm{SL}(2,\mathbb{R})$ connections [4, 121, 20, 28, 122].

To quantize the reduced phase space, $\widetilde{P} = T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$, which is manifestly a homogeneous space for $\mathit{Diff}^+(S^1)$, we employed Isham's group-theoretic approach to canonical quantization. From the structure of this phase space, we found that a natural canonical group on which to base the quantization is $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$.

When realized as symplectic symmetries of the phase space, the group interestingly reduces to BMS_3 (i.e., the central charge associated with *Vira* is trivially realized, matching with the algebra obtained in [43]). The quantum theory is thus constructed to carry an irreducible unitary representation of the associated algebra, $\mathfrak{bms}_3$, which according to Mackey's theory of induced representations is characterized by a choice of a coadjoint orbit of Virasoro together with a choice of a (projective) irreducible unitary representation of the corresponding little group [45]. In the most natural case, according to the Casimir matching principle, the quantum states are identified with wavefunctions "living" on $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ with an "internal index" in some projective irreducible unitary representations of $PSL(2,\mathbb{R})$ (or, assuming CMC time reversal symmetry, a true irreducible unitary representation of $\mathrm{SL}(2,\mathbb{R})$). This canonical group appeared naturally in our reduction procedure, but it is worth noticing that there is a non-trivial symplectomorphism between $T^*\mathcal{Q}$, with the natural symplectic form associated with its cotangent bundle structure, and $\mathcal{Q}\times\mathcal{Q}$, with the natural symplectic structure that each factor $\mathcal{Q}$ inherits as a coadjoint orbit of Virasoro [31]. Thus, while those two phase spaces are isomorphic, the quantization for the latter would most naturally be based on the group $\mathit{Vira}\times\mathit{Vira}$, thus producing a different set of representations. In fact, $\mathcal{Q}\times\mathcal{Q}$ is the most natural realization for the reduced phase space of asymptotically AdS_3 spacetimes, and $\mathit{Vira}\times\mathit{Vira}$ is the symmetry structure of a CFT_2, which is a manifestation of the holographic principle [30, 29, 28].

We emphasize that we have only constructed a non-perturbative *kinematical* quantization of the system, i.e., obtained a family of possible quantum theories that are naturally associated with the classical phase space, carrying a representation of the canonical observables. A complete quantization needs in addition a description of the quantum dynamics, i.e., finding a suitable manner to represent the classical time-evolution Hamiltonian as a self-adjoint operator on the Hilbert space. While we have a classical description of the Hamiltonian generating evolution in CMC time, whose value has the very simple geometric interpretation of measuring the area of the CMC slice [40], it is expressed as a highly complicated formula in terms of the reduced phase space variables. In particular, that formula involves the solution to the Lichnerowicz equation, which cannot be written in closed-form. We do not see a straightforward way to effectuate such a quantization, at least non-perturbatively, for even if an explicit formula for the Hamiltonian as a function of the canonical charges were to be found, there would likely be serious operator-ordering ambiguities which would need to be resolved. A more realistic goal would be to address the dynamical portion of the quantization at the perturbative level. As we have seen in Sec. 7.2, there are regimes where the Hamiltonian can be approximated, even becoming "free" (i.e., quadratic in momentum-like variables) for states sufficiently close to the symmetric diamond, in the limit of large boundary length compared to the AdS length. Notice, however, that the symmetric diamond is not a "classical vacuum state" for the time-evolution Hamiltonian since there are states with lower "energy" (i.e., area), even in a neighborhood of that state. In fact, while the classical Hamiltonian is manifestly bounded from below, it has no minimum, much like the Hamiltonian for Liouville field theory whose potential has an exponential form $V(\phi) = e^{2b\phi}$ (indeed Liouville theory has been linked to three dimensional gravity before [73, 74, 75, 76, 77, 78, 79]).

Another set of questions concerns what a "quantum causal diamond" really means. At the classical level we arrived at the interpretation that the states (i.e., points in the reduced phase space) correspond to shapes of causal diamonds embedded into AdS_3 (or $Mink_3$ if the cosmological constant vanishes). That is, a causal diamond with a round maximal slice and another one with an oval maximal slice are distinct classical states, with physically distinguishable properties (e.g., if one shoots geodesics across the maximal slice, starting normal to the boundary at one point and ending at the other side of the boundary, in the first case they would all have the same length while in the second case some would be longer than others). In fact, we have shown that there is a map from each point of the reduced phase space to a causal diamond in AdS_3 or $Mink_3$ (or, more simply, to a loop of length ℓ that is the boundary of a spacelike disc). However, for this interpretation to hold it is essential that both the Q_n *and* P_n coordinates are specified simultaneously, which of course is not feasible in the quantum theory since these observables do not commute. This apparent breakdown of the fabric of spacetime is expected in a generic theory of quantum gravity, and there is no reason to believe that a quantum spacetime should have properties similar to those of a manifold (except of course in some semi-classical limit). In our specific realization, in addition to the non-commutativity of the geometrical variables, note also that: as the CMC time is chosen before quantization, it retains a classical character in the quantum theory, thus standing on a different footing than "space" in the quantum picture; second, even the concept of a spatial geometry is deformed since, while the configuration variables (i.e., the $\widehat{Q}_n$'s) do commute, they are not in direct correspondence with the spatial geometry of the CMC slices, but rather with the *conformal* geometry of the slices. Accordingly, it would be valuable to better understand what kind of "spacetime" emerges, from our quantization, as the "interior" of a quantum causal diamond, possibly shedding some light on what is the meaning of a "subsystem of a quantum spacetime", or even what is a "quantum spacetime" itself.

To make sense of the quantum geometry of causal diamonds, a reasonable first step is to understand the physical meaning of the canonical observables. In this paper we mostly focused on the meaning of the momentum charges (i.e., the P_n's), which were found to be related to Fourier modes of the component $K_{ab}t^a n^b$ of the extrinsic curvature of the CMC slices at the corner, where t^a and n^b are unit vectors on the slice tangent and normal, respectively, to the boundary. Of distinguished significance is P_0 which can be interpreted as the twist of the corner loop, therefore being classically a property of the diamond shape itself (i.e., independent of choices of spatial slices). In the quantum theory we showed that the twist is quantized, in any non-trivial representation, in integer (or half-integer) multiples of $16\pi^2\ell_P/\ell$. This is a non-trivial result since the twist is a continuous parameter classically (in fact, we can see that when the length of the loop is much larger than the Planck length, the spacing between twist levels goes to zero, recovering the expected classical behavior). One might speculate that there is some sense in which this result can extrapolated to a general statement about the twist of loops in three-dimensional gravity (and perhaps even in the presence of matter fields, since the "twist quantum" is independent of the cosmological constant).

The physical interpretation of the configuration charges (i.e., the Q_n's) has proved much more elusive. Although we have given explicit formulas for these charges in terms of the $PSL(2,\mathbb{R})$-class of the boundary diffeomorphism, which in turn is explicitly related to a conformal class of spatial geometries on the CMC slices, we could not provide an interpretation in terms of simple geometrical properties of the shape of the causal diamond. A path currently under consideration attempts to understand the structure of corner symmetries directly at the ADM level, by analyzing the differentiability conditions for the ADM Hamiltonians (i.e., the boundary terms that need to be included so that the corresponding Hamiltonians generate regular symplectic flows on the phase space, and consequently have well-defined Poisson brackets with other charges). We find that the P_n's are precisely the charges associated with diffeomorphisms tangent to the corner, but there are another two sets of charges associated with deformations of the corner in the normal directions. These normal deformation charges, when evaluated on the constraint surface, are related to the expansion parameter of the null rays of the future and past horizons at the corner. In this paper we have shown that two families of charges (P's and Q's) should suffice to parametrize the reduced phase space, so it may be speculated that the Q_n's are related to some function of the expansion parameters of the null generators of the horizons. Another hint in this direction is that BMS_3 is, as we have mentioned, notably associated with asymptotic symmetries on the null infinity of asymptotically-flat spacetimes in three dimensions. It is thus plausible that there exists an alternative point of view in which the causal diamonds are described by charges directly associated with the horizons in a neighborhood of the corner, making the analogy with asymptotically-flat spacetimes more explicit.

Finally, we comment that it would be interesting to investigate the thermodynamic and entanglement properties of the quantum causal diamonds. A Holy Grail of quantum gravity is still to explain the Bekenstein-Hawking entropy formula from a microscopic point of view, and this entropy is widely believed to correspond to a measure of the entanglement across the black hole horizon. In fact, it is expected that a similar formula for the entanglement entropy should also hold for generic entangling surfaces. Naturally this can be analyzed in the context of causal diamonds, e.g., [123, 124, 37, 125, 126, 36, 127, 128, 129, 130, 131]. The research on the thermodynamics of causal diamonds has been mainly perturbative in nature, so having a concrete quantum theory of causal diamonds could be valuable in understanding their thermodynamic and entanglement structures from a non-perturbative perspective. In the microcanonical ensemble, the entropy is given by the logarithm of the number of states compatible with specified values of some "macroscopic" charges, typically the energy and angular momentum. Taking the "energy" to mean the CMC time-evolution Hamiltonian, the entropy $\mathcal{S}(\mathcal{E},\mathcal{J})$ should be defined as the logarithm of the number of eigenstates of H and P_0 with eigenvalues (around), respectively, $\mathcal{E}$ and $\mathcal{J}$.[70] Curiously, in the entanglement equilibrium argument [123, 37], variations of the maximal slice with fixed volume play a key role — this condition is quite analogous, in our case, to fixing the value of H (at $\tau = 0$), as it has the interpretation of being the "volume" (in two dimensions) of the maximal slice. To properly evaluate the dimension of the subspace with $H \leq \mathcal{E}$ requires us to understand its spectrum, which requires us to understand how to represent H in the Hilbert space. However, as mentioned earlier, this is a topic for future research.

[70]One could also consider the entropy associated with fixing Q_0 (and P_0), as Q_0 can be regarded as some kind of "quasi-local energy". However, the spectrum of Q_0 is continuous and thus it seems that a corresponding entropy can not be properly defined.

Part III

Perspectives and outlooks

To conclude this thesis, we will discuss some general aspects surrounding the project, including possible ramifications and other ventures motivated by it, reaching beyond the main line of this work. In particular, we have in mind here quantum gravity in a broader perspective, not necessarily limited to (2+1)-dimensional causal diamonds. In this section, a higher level of speculation and personal philosophical inclinations are expressed, and we include ideas that are unpolished and consequently not on the same standard of rigor as the rest of the thesis.

14 On quantum mechanics (and canonical quantization)

As explained in the introduction, our starting premise was based on two principles: (*i*) that quantum gravity is a quantum mechanical theory of gravity and (*ii*) that this quantum theory is obtained from a canonical quantization of general relativity. While this may be a perfectly reasonable axiomatic basis, it could be worthwhile to investigate deeper the roots of these assumptions.

On point (*i*), it is not difficult to imagine that a theory describing physics at the Planck scale may look nothing like quantum mechanics, but only reduce to quantum mechanics and general relativity in their appropriate regimes. A humbling exercise is to recall that the jump in length scales from "classical physics" to "atomic physics" is roughly 7 orders of magnitude (i.e., from small quotidian scales, $\sim 1\,\mathrm{mm}$, to the radius of a hydrogen atom, $\sim 10^{-10}\,\mathrm{m}$), whereas the conceptual jump from Newtonian physics to quantum mechanics is immense; while it is quite remarkable that quantum mechanics works so well (in its relativistic field-theoretical incarnation) all the way down to the minuscule scales accessible by the Large Hadron Collider (i.e., $\sim 10^{-19}\,\mathrm{m}$ or $\sim 10\,\mathrm{TeV}$), this is still 16 orders of magnitude away from the Planck scale (i.e., $\sim 10^{-35}\,\mathrm{m}$). Thus, we may well be much further from blazing a trail towards quantum gravity than those at Newton's time would have been in anticipating quantum mechanics. Alternatively, a more optimistic view would be that, as now the knowledge of physics is much more vast and robust, the constraints in developing novel theories, in a way which is consistent with all current observations, are much tighter. In any case, it is important to reflect on whether a satisfactory theory of "quantum gravity" has not been yet achieved merely because of the technical complications and lack of clarity in the efforts to embed gravity into the framework of quantum mechanics, or rather because of some fundamental incompatibility between these two domains.

The dynamical nature of the spacetime is certainly an aspect of gravity that poses some deep conceptual challenges. For example, in the standard conception of physics, it appears that all kinds of measurements ultimately refers to observing positions and times, and the "irreducible components" of any measuring device are rulers and clocks. In traditional quantum mechanics (and quantum field theory), the spacetime is a classical entity on which the concept of *events* have a natural objective reality — moreover, the spacetime appears from our experience to be "smooth", in the sense that it can be appropriately modeled as a real differentiable manifold. When the very notion of spacetime is assumed to become quantum, the associated notion of events must be reinterpreted, raising the question of "*what an observation really is?*" It has been argued [132, 133, 134] that even the motivation for using real (or complex) numbers in quantum gravity may be questionable: if observations always come down to measuring distances and times, and one admits that spacetime may fundamentally be modeled on something different than a real differentiable manifold, then why should we expect physical quantities to be real (or complex) valued? This questioning may motivate one to look for alternative categorical frameworks for physics, e.g., theories constructed directly on a *Topos* different from the category of *Sets*.[71] Another issue concerns the concept of probability, which is central to quantum mechanics, but becomes unclear if one understands quantum gravity to include cosmology: if the universe is the "whole", it cannot be replicated to allow a frequentist interpretation of probabilities; and if there is a single state, what would be the need for a Hilbert space?

It should be noted that these concerns refer to the ambitious attempt of formulating a fundamental theory of quantum gravity, or at least one that brings us

[71] *Topoi* are a special kind of category in which sub-objects (of any object) have a (pseudo-Boolean) logical structure. *Sets* is a special *Topos* whose underlying logic is Boolean. This ingrained logical structure offers a general framework to write any theory of physics in a *realist sense*, i.e., so that one can associate a logical algebra to propositions about the system, which have "truth-values" when evaluated on "states". The requirement of realism seems particularly desirable if the goal is to construct a theory of the whole universe, without external observers.

closer to describing physics near the Planck scale. There is nothing, as far as we know, inconsistent with the effective, perturbative incorporation of gravity into the framework of quantum field theory, as long as one is content with describing physics at low energies [135]. In fact, the effect of gravitons (or quantized metric perturbations) is consistent with cosmological predictions (e.g., on the cosmic microwave background radiation) in simple models of inflation [136]. Lingering skepticism may also be (at least partially) addressed with the recent proposals of "tabletop experiments" for quantum gravity [137], which are expected to provide further evidence of the (standard) quantum nature of gravitons in the near future.

On point (*ii*), it seems desirable to better understand what is really behind canonical quantization. In Sec. 9.1, we explained that one *justification of principle* is that dynamics is expressed in very analogous ways in the classical (Poisson) and quantum (commutator) algebras, offering a natural manner to construct a quantization map that is *preserved in time*. But the concept of time in gravity is much different than in other domains where canonical quantization has been successfully applied: in gravity time is dynamical, or better, there is no preferred "time" as everything is relational.[72] It is also worth noticing that, at the classical level, the only observational content of a theory is the *space of solutions* to the equations of motion of the observables, which only determines the *topology* of the (reduced)

[72] In standard quantum field theory there is also no single preferred "time", but there is a preferred "class of times" defined by inertial frames of reference. Even in the context of field theory in curved spacetimes, any time function which defines a foliation of the spacetime by Cauchy surfaces is sufficient to construct a consistent quantum theory. Of greater importance here is that the notion of causality is fixed (non-dynamical), while in gravity even the causal structure is dynamical and state-dependent.

phase space. The symplectic structure is something extra: it depends on the form of the equations of motion (or a choice of Hamiltonian).[73] We may thus worry that canonical quantization is, at its roots, too intertwined with a notion of an external, non-dynamical time. Adding to this, in the typical application of canonical quantization, the (pre-)phase space is parameterized by "initial data" on a slice of time, which is another reference to classical notion of time. We will return to this point when we discuss the CMC gauge-fixing of time (Sec. 17).

To be clear, the concerns enunciated above refer only to an "a priori justification" of canonical quantization, and it is entirely possible that it could still be valid in its pure form, when abstracted to pose that there is a fundamental connection between certain classical (Poisson) and quantum (commutator) algebras of observables, without explicit reference to a time-evolution. It would be interesting to identify general characteristics or conditions in classical and quantum theories expected to be necessary for canonical quantization to work. Say, given a quantum theory $\mathcal{T}$ with classical limit $\mathit{classical}(\mathcal{T};\mathit{conditions})$, where "*conditions*" specifies how this limit is defined (e.g., some energy restriction, decoherence, coarse-graining, reinterpretation, renormalization, etc), then is there a quantization such that $\mathit{quantization}(\mathit{classical}(\mathcal{T};\mathit{conditions})) = \mathcal{T}$? As it is, the question is too vague, but one could start with more specific, simple examples. In particular, one could consider situations where the classical theory has emergent features not present in

[73]Relatedly, two actions are classically equivalent provided that they are stationary at the same configuration histories, but the symplectic form also depends on how the action behaves in a neighborhood of its stationary points.

the original quantum theory, and ask questions such as "is QCD hidden among the possible quantizations of the chiral model?" or "could one infer that matter is made up of particles from quantizations of hydrodynamics?".

We will return to non-linear sigma models in Sec. 15.3, but let us briefly comment on the case of hydrodynamics here. At the deeper lever, a fluid consists of a collection of quantum particles, and due to the discrete (point-like) nature of these particles there is no explicit diffeomorphism symmetry. At the classical coarse-grained level, this bunch of (decohered) particles appear to form a continuum, "the fluid", and all fluid configurations (say, described by the velocity flow) are related through diffeomorphisms. In the quantization of this fluid, it is thus natural to consider unitary representations of the diffeomorphism group. But how accurate is the resulting picture? Of course, many representations will fundamentally retain this "continuous aspect" (e.g., wave functions on the classical configuration space), but what is interesting[74] is that there are also representations associated with discrete "configuration spaces" [138]. While the actual dynamics of these "pointlike quantum objects" may not be entirely accurate, one may at least take it as a hint of what "could be" making up the fluid. As we have mentioned, in case of our causal diamonds, the canonical group was actually $(\mathfrak{vira}^*)^* \rtimes \mathit{Vira}$, and just for simplicity we chose to assume $(\mathfrak{vira}^*)^* \sim \mathfrak{vira}$; but it should be stressed that one could consider more exotic "distributional" group actions, and also representations based on non-regular orbits, at least to get a flavor of what "could be" those more fundamental

[74]Mentioned to me by Laurent Freidel.

structures making up the spacetime.

While it is healthy to keep such points in mind, it is also clear that without strong physical intuition, and possibly new experimental inputs, it would be easy to fall into misguided philosophical ventures. This perspective is what motivates our conservative approach, as expressed (in a much summarized form) in the introduction, Sec. 1. That is, while we cannot justify *a priori* the validity of quantum mechanics and the applicability of canonical quantization, it is valuable to understand exactly what it *can* and what it *cannot* give us. In this work, we tried to honor this perspective by being as careful as possible, both with the non-perturbative treatment of the constraints of the underlying theory (general relativity) and with the application of a method of canonical quantization that respects the global structure of the phase space.

15 On non-trivial phase spaces

One aspect that we have greatly emphasized is that the quantization scheme should respect the global structure of the phase space. Classical mechanics is very local in phase space, in the sense that if we only know about a neighborhood $\mathcal{V}$ of a point $p \in \mathcal{P}$ of the phase space, then we can still flow, from that point, with respect to any Hamiltonian, until we reach the edge of $\mathcal{V}$. Thus, the global structure of the phase space is only noticed once a trajectory runs for enough time so as to "wrap around" the phase space. In quantum mechanics, states are fundamentally "spread out", and the global structure of the phase space often has deep influence in the quantum theory, like affecting how operators are represented on the Hilbert space, their spectrum and their properties (such as self-adjointness, unitarity, etc).

If one wishes to investigate questions such as the fundamental structure of the spacetime, it is important to consider these details. For example, one of the few available and widely accepted theoretical tests for a quantum gravity theory is to explain the black hole entropy from a microstate counting, but one would expect that such a counting would be sensitive to the spectrum of basic observables defining the black hole (e.g., its mass and spin).

A very familiar example where the topology of the configuration space affects the spectrum of certain operators is the case of a particle on a circle ($\mathcal{P} = T^*S^1 = S^1 \times \mathbb{R}$): the momentum does not have a continuous real spectrum, but rather it is discretized in integer multiples of h/ℓ, where ℓ is the length of the circle. Next we discuss some other examples of non-trivial phase spaces that we have explored.

15.1 A particle on the sphere

The case of a particle on a 2-sphere is possibly one of the simplest where the phase space is not a just obtained by (linearly) cutting and gluing an Euclidean space. Together with Ted Jacobson, I analyzed this case, including a magnetic monopole flux, as a warm-up exercise to better understand Isham's group theoretic quantization [47]. The abstract is replicated below:

"*The problem of quantizing a particle on a 2-sphere has been treated by numer-*

ous approaches, including Isham's global method based on unitary representations of a symplectic symmetry group that acts transitively on the phase space. Here we reconsider this simple model using Isham's scheme, enriched by a magnetic flux through the sphere via a modification of the symplectic form. To maintain complete generality we construct the Hilbert space directly from the symmetry algebra, which is manifestly gauge-invariant, using ladder operators. In this way, we recover algebraically the complete classification of quantizations, and the corresponding energy spectra for the particle. The famous Dirac quantization condition for the monopole charge follows from the requirement that the classical and quantum Casimir invariants match. In an appendix we explain the relation between this approach and the more common one that assumes from the outset a Hilbert space of wave functions that are sections of a nontrivial line bundle over the sphere, and show how the Casimir invariants of the algebra determine the bundle topology."

Notably, see how even some background parameters, which are not operators in the theory, may have their value constrained by some matching principle. In particular, here the famous Dirac charge quantization condition (i.e., $eg = n\hbar/2$, where e is the electric charge, g is the magnetic charge and $n \in \mathbb{Z}$) follows from the Casimir matching principle together with the assumption that the quantum theory is non-trivial.

It is also interesting to see how, in this group-theoretic approach, the exact quantization compares with a "perturbative quantization", where one simply uses local phase space conjugate coordinates, (X_1, P_1, X_2, P_2), as the basis of the quantization. Consider the excerpt below, from the same paper:[75]

"*The difference between the quantizations on a plane and on a sphere can be understood from a group-theoretic perspective, in terms of the "planar limit" of the sphere, as follows. We expect that a particle that remains near the north pole of the sphere at all times should not be able to "feel" the global structure of the sphere. Thus, in some limit, the quantum mechanics on a sphere must reduce to the usual one on a plane. To see how this works, consider a "sector" of the Hilbert space in which* $X_1 := N_1 \sim \theta$, $X_2 := N_2 \sim \theta$ *and* $N_3 \approx 1$, *where* $\theta \ll 1$ *is the angle around the north pole where the particle is localized. Note that the "vertical momentum"* $P_3 \sim \theta|P|$ *is small in this sector, so we can approximate* $J_1 \approx -P_2 - egX_1$, $J_2 \approx P_1 - egX_2$ *and* $J_3 \approx X_1P_2 - X_2P_1 - eg$, *where terms of order* θ^2 *were neglected. In this sector, the algebra reduces to* $[X_i, P_j] = i\hbar\delta_{ij}N_3$, J_3 *behaves as the generator of rotations for* X_i *and* P_i, *and* N_3 *becomes a central element (taking the value 1 in the relevant representation). This deformation of the algebra is known as the Inönü-Wigner contraction. At the group level, this corresponds to a deformation of the Euclidean group* $E_3 = \mathbb{R}^3 \rtimes SO(3)$ *into* $(\mathbb{R}^2 \times \mathbb{R}) \rtimes (\mathbb{R}^2 \rtimes SO(2))$, *where the first factor is generated by* $(X_1, X_2; N_3)$ *and the second by* $(P_1, P_2; J_3)$. *This contracted group can be reexpressed as* $H(2) \rtimes SO(2)$, *where* $H(2)$ *is the Heisenberg group in*

[75] *Notation*: consider that the Euclidean group, $E(3) = \mathbb{R}^3 \rtimes SO(3)$, is generated by the charges $(N_1, N_2, N_3; J_1, J_2, J_3)$, where N_i are the components of a unit vector specifying the angular position of the particle and J_i are the components of angular momentum (modified by the magnetic flux, $\vec{J} = \vec{N} \times \vec{P} - eg\vec{N}$).

two spatial dimensions, generated by $(X_1, P_1, X_2, P_2; N_3)$*, and* $SO(2)$ *is the rotation group around the origin, generated by* J_3*. Note that* $S := J_3 - (X_1P_2 - X_2P_1) = -eg$ *is a Casimir operator, interpreted as the intrinsic spin of the particle. Since* J_3 *differs from* $X_1P_2 - X_2P_1$ *only by a Casimir operator, it follows that the irreducible representations of* $H(2) \rtimes SO(2)$ *are also irreducible when restricted to* $H(2)$*. As* $H(2)$ *has a unique irreducible unitary representation, this confirms that we do in fact recover the quantum mechanics on a plane (for any value of the intrinsic spin Casimir* S*). It is interesting to note an important difference between the plane and the sphere: the subgroup* $SO(2)$ *of* E_3 *was "pulled out" of* $SO(3)$ *during the deformation, so it appears as a factor in* $H(2) \rtimes SO(2)$ *rather than as a subgroup of* $SO(3)$*. The spin is therefore not quantized on a plane, because the* $SO(2)$ *gets "unwrapped" to* $\mathbb{R}$ *when considering projective representations (that is, when considering the universal cover of the group). Thus the quantization of the spin on a sphere is a truly topological effect."*

15.2 A particle on the half-line

Another case that I have studied (with a paper in preparation [139]) is the quantization and quantum mechanics of a particle living on the half-line, $\mathbb{R}^+$. Such a dynamical system is associated with a phase space that is the cotangent bundle of the configuration space, $T^*\mathbb{R}^+ = \mathbb{R}^+ \times \mathbb{R}$, and the symplectic form is given by $\omega = dp \wedge dx$, where $x \in \mathbb{R}^+$ is a position (configuration) coordinate and $p \in \mathbb{R}$ is the conjugate momentum.[76] The half-line is to be regarded as the entire physical space, instead of a subspace of the line, so an intrinsic perspective is imperative. One can think of this system as describing certain mini-superspace models of cosmology [140, 141, 142, 143], where the configuration variable is the volume of the universe (and thus strictly positive) and the conjugate variable is the Hubble constant; or (1+1)-dimensional gravity from the reduced phase space point of view [144], for the only gauge-invariant property of a metric on a 1-dimensional space is the total proper length, another strictly positive configuration variable.

As emphasized earlier, the traditional Dirac quantization method is not appropriate since there is no unitary irreducible representation of the Heisenberg algebra in which $\widehat{x}$ has a spectrum contained in $\mathbb{R}^+$. A natural alternative is the *affine quantization*, which has been advocated by many authors [145, 146, 147, 148, 149, 92, 93]. The focus of my analysis is on how this different quantization affects the self-adjointness of certain operators, including a class of simple time-evolution Hamiltonians. It is a well-known mathematical fact that, in infinite-dimensional Hilbert spaces, being symmetric is a necessary but not sufficient condition for being self-adjoint, which means that it is possible to have non-unitary dynamics (in particular, with the possibility of "leaking away" from the configuration space) even if the Hamiltonian is symmetric.[77] In particular, the choice of operator-orderings, within a

[76] Note that the phase space is topologically trivial, in the sense of being symplectomorphic to $\mathbb{R}^2$, with conjugate coordinates $q := \ln x$ and $\pi := xp$. This is somewhat analogous to the case of the diamond, where the phase space is also contractible (see App. D).

[77] A dense (possibly unbounded) linear operator A is *symmetric* if $A^\dagger = A$ in $\mathrm{Dom}(A^\dagger) \cap \mathrm{Dom}(A)$, and it is *self-adjoint* if additionally $\mathrm{Dom}(A^\dagger) = \mathrm{Dom}(A)$.

class motivated by the affine symmetry, is crucial in determining the self-adjointness of the Hamiltonian.

Curiously, it is revealed that even a classically "free" Hamiltonian, $H = p^2/2m$, which certainly does not lead to a classically closed system (i.e., there is nothing preventing the particle from reaching $x = 0$ in finite time), can still be made self-adjoint for certain operator-orderings, thus making the system closed quantum mechanically. Morally, this phenomenon can be seen as the *opposite of tunneling*, in that quantum effects would *help* closing the system. We also compare the affine quantization of a particle on the half-line with the Dirac quantization of a particle on the line, which is confined to the half-line by a potential barrier, showing that there are subtle differences in the conditions for unitarity of time evolution.

Perhaps even more surprisingly, we discover a phenomenon where a potential that diverges to $-\infty$ as $x \to 0$ *may* lead to a self-adjoint Hamiltonian $H = p^2/2m + V(x)$. In such a case, the time evolution is unitary and thus the particle is dynamically confined to the half-line. In other words, it is as if the potential acts as a wall, even though it is really a precipice, holding the particle in the "positive half" of the line. This is highly non-intuitive from a classical perspective since such a potential attracts the particle to $x = 0$, in an accelerated fashion, thus pushing it towards the "negative half" of the line in finite time (in fact, the classical system is only closed if the potential diverges to $+\infty$). The explicit examples of such exotic potentials are constructed in a particular way where they diverge to $-\infty$ as $x \to 0$ in a certain oscillatory manner. For this reason, we may wish to call this phenomenon "resonance-induced barricade".

15.3 Non-linear sigma model

Another interesting problem, of greater ambition, is to develop a fully non-perturbative quantization of the non-linear sigma model. The common analysis of non-linear sigma models is usually done in a perturbative framework, looking only at a small neighborhood of a point in the target space and thus only probing the "local quantization". (See related discussion in Sec. 15.1.) In a paper being prepared by myself [150], I attempt to "set up" the problem for global quantization by describing a transitive symmetry group that acts naturally on the phase space. As this group (likely) admits many unitary irreducible (projective) representations, it would be interesting to see what kinds of quantum theories lie among the possibilities. In particular, as certain low-energy limits of QCD can be described by non-linear sigma models (namely, the chiral model), one may wonder how much of QCD could be "inferred" just from this effective model. This exercise might teach us about our chances of inferring something fundamental about quantum gravity just from general relativity.

I present here a few formulas for the non-linear sigma model whose configuration space is $\mathcal{C} = \mathrm{Maps}(\mathbb{R}^3, S^2)$, that is, where the spacetime manifold is $\mathbb{R} \times \mathbb{R}^3$ and the target space is S^2. The action is

$$S = \frac{1}{2} \int d^4x \sqrt{-g}\, g^{\mu\nu} \kappa_{AB} \partial_\mu \phi^A \partial_\nu \phi^B$$

where ϕ^A are local coordinates on S^2 and κ_{AB} is the $SO(3)$-invariant metric on it (normalized so that the sphere has area 4π). In Cartesian coordinates for the spacetime, the corresponding Hamiltonian is

$$H = \frac{1}{2}\int d^3x \left(\kappa^{AB}\pi_A\pi_B + \kappa_{AB}\nabla\phi^A \cdot \nabla\phi^B\right)$$

where π_A are the conjugate momenta.

A natural canonical group would be one that acts (on field variables) as a 3-dimensional Euclidean group at each point in (physical) space,

$$G = \mathrm{Maps}(\mathbb{R}^3, \mathbb{R}^{3*} \rtimes SO(3)) = \mathrm{Maps}(\mathbb{R}^3, \mathbb{R}^{3*}) \rtimes \mathrm{Maps}(\mathbb{R}^3, SO(3))$$

A generic group element is denoted by (α, R), and a generic algebra element is denoted by (α, η). Let us think of the configuration space as the orbit of $\mathrm{Maps}(\mathbb{R}^3, SO(3))$ acting on a given vector $\chi_0 \in \mathrm{Maps}(\mathbb{R}^3, \mathbb{R}^3)$, say $\chi_0(x) = e_3$ (i.e., north-pole in target space, constant in space). In components, the field will be denoted by $\chi^i(x)$, where $i = 1, 2, 3$ is an index in the "embedding target space", satisfying $\chi^i\chi^i = 1$.

The canonical charges are given by

$$P(\eta) = \int d^3x\, \pi(X_\eta)$$
$$Q(\alpha) = \int d^3x\, \alpha(\chi)$$

where $X_\eta(x)$ is the vector field on the sphere (based at the spatial point x) induced by the rotation generator $\eta(x)$. In components (in the embedding target space), we have $\pi(X_\eta) = \pi_i X^i_\eta$ and $\alpha(\chi) = \alpha_i \chi^i$. In terms of these charges, the Hamiltonian reads

$$H = \frac{1}{2}\int d^3x \left(P_i(x)P_i(x) + \nabla Q^i(x) \cdot \nabla Q^i(x)\right)$$

in which

$$P_i(x) := P(e_i \delta_x)$$
$$Q^i(x) := Q(e^i \delta_x)$$

where $\delta_x(x') = \delta(x' - x)$ is the Dirac delta function. These operators satisfy the algebra

$$[P_i(x), P_j(x')] = i\epsilon_{ijl}\delta(x - x')P_l(x)$$
$$[P_i(x), Q^j(x')] = i\epsilon_{ijl}\delta(x - x')Q^l(x)$$
$$[Q^i(x), Q^j(x')] = 0$$

with the Casimir conditions

$$Q^i(x)Q^i(x) = 1$$
$$P_i(x)Q^i(x) = 0$$

Using Mackey's theory, and assuming the trivial representation for the little group $H = \text{Maps}(\mathbb{R}^3, U(1))$, we get a wave function representation, $\Psi : \text{Maps}(\mathbb{R}^3, S^2) \to \mathbb{C}$, for this algebra

$$(P_i(x)\Psi)(\chi) = -i\epsilon_{ijl}\chi^j \frac{\delta\Psi}{\delta\chi^l(x)}$$

$$(Q^i(x)\Psi)(\chi) = \chi^i(x)\Psi(\chi)$$

where Ψ is taken as a function on $\text{Maps}(\mathbb{R}^3, \mathbb{R}^3)$. Note that the expressions above are independent on how Ψ is defined away from the sphere $\chi^i\chi^i = 1$. In this representation, the Hamiltonian acts as

$$H\Psi(\chi) = \frac{1}{2}\int d^3x \left[-\frac{\delta}{\delta\chi^i}\left((|\chi|^2\delta_{ij} - \chi_i\chi_j)\frac{\delta\Psi}{\delta\chi^j} \right) + \nabla\chi^i \cdot \nabla\chi^i \, \Psi(\chi) \right]$$

Notice the appearance of the projector inside the derivative, ensuring that the result will not depend on how Ψ is defined away from the sphere.

We can also work in momentum space, which is convenient because of the spatial derivatives. Define,

$$\widetilde{P}_i(k) := P(e_i\rho_k) = \int d^3x \, e^{ikx} P_i(x)$$

$$\widetilde{Q}^i(k) := Q(e^i\rho_k) = \int d^3x \, e^{ikx} Q^i(x)$$

where $\rho_k(x) = e^{ikx}$ is the plane wave function. These are just the Fourier transform of the position-based charges. They satisfy the algebra

$$[\widetilde{P}_i(k), \widetilde{P}_j^\dagger(k')] = i\epsilon_{ijl}\widetilde{P}_l(k-k')$$
$$[\widetilde{P}_i(k), \widetilde{Q}^{j\dagger}(k')] = i\epsilon_{ijl}\widetilde{Q}^l(k-k')$$
$$[\widetilde{Q}^i(k), \widetilde{Q}^{j\dagger}(k')] = 0$$

and the reality conditions $\widetilde{P}_i^\dagger(k) = \widetilde{P}_i(-k)$ and $\widetilde{Q}_i^\dagger(k) = \widetilde{Q}_i(-k)$. In terms of these, the Hamiltonian reads

$$H = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\left(\widetilde{P}_i^\dagger(k)\widetilde{P}_i(k) + |k|^2\widetilde{Q}^{i\dagger}(k)\widetilde{Q}^i(k)\right)$$

We can also define "a-operators" as

$$a_i(k) := \widetilde{P}_i(k) + i|k|\widetilde{Q}^i(k)$$

so that the Hamiltonian becomes

$$H = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3} a_i^\dagger(k) a_i(k)$$

The algebra of these operators are not so nice though,

$$
\begin{aligned}
[a_i(k), a_j^\dagger(k')] &= i\epsilon_{ijl}\left(\frac{|k-k'|+|k|-|k'|}{2|k-k'|}\,a_l(k-k') + \frac{|k-k'|-|k|+|k'|}{2|k-k'|}\,a_l^\dagger(-k+k')\right)\\
[a_i(k), a_j(k')] &= i\epsilon_{ijl}\left(\frac{|k+k'|+|k|+|k'|}{2|k+k'|}\,a_l(k+k') + \frac{|k+k'|-|k|-|k'|}{2|k+k'|}\,a_l^\dagger(-k-k')\right)\\
[a_i^\dagger(k), a_j^\dagger(k')] &= -i\epsilon_{ijl}\left(\frac{|k+k'|+|k|+|k'|}{2|k+k'|}\,a_l^\dagger(k+k') + \frac{|k+k'|-|k|-|k'|}{2|k+k'|}\,a_l(-k-k')\right)
\end{aligned}
$$

We see that although the Hamiltonian looks very simple (in fact, it looks "free"), all the "non-linearity" of the model is encoded in the algebra of operators. In particular, notice how the algebra (in momentum space) mixes different spatial scales.

The quantization *per se* will be left for future work, but it would involve finding unitary irreducible representations of this algebra which also carries a representation of the Poincaré group. Note that, in such a representation, it would automatically follow from the "free" form of the Hamiltonian that its spectrum is bounded from below.

16 On subsystems

The notions of subsystems or subregions are quite fuzzy in quantum gravity, possibly much more than the notion of spacetime itself.

Let us start with the naive picture of a "fluctuating geometry" on a fixed manifold, and try to think of gravity as typical QFT. Since spacetime diffeomorphisms are gauge transformations in gravity, there can be no (gauge-invariant) observables that are supported in a (compact, proper) subregion of the spacetime. That is, say an observable $\mathcal{O}$ is supported in a (not necessarily proper) region $\mathcal{V}$ of the (connected) spacetime $\mathcal{M}$, in the sense that $\mathcal{O}$ is constructed out of local fields $\phi(x)$ (including the metric) with $x \in \mathcal{V}$. In a suggestive notation, we may write $\mathcal{O} = F[\phi(x)]$ (but note that F has infinitely many arguments, indexed by $\mathcal{V}\times$ *"number of fields"*). Now consider a gauge transformation which acts on local fields as $\Psi\phi := \psi^*\phi$, where $\psi \in$ *Diff*$(\mathcal{M})$, and consequently acts on $\mathcal{O}$ as

$$\Psi^*\mathcal{O} = \Psi^* F[\phi(x)] = F[\Psi\phi(x)] = F[\psi^*\phi(x)]$$

Since $\psi^*\phi(x)$ is supported at $\psi(x)$, $\Psi^*\mathcal{O}$ is thus is supported in $\psi(\mathcal{V})$. Now, if $\mathcal{O}$ is gauge-invariant, it must be that $\mathcal{O} = \Psi^*\mathcal{O}$ for all Ψ, and in particular

$$\mathcal{V} = \psi(\mathcal{V})$$

for all ψ. This is only possible if: (i) $\mathcal{V} = \emptyset$; (ii) $\mathcal{V} = \mathcal{M}$; or (iii) $\mathcal{V} = \partial\mathcal{M}$ (where $\partial\mathcal{M}$ denotes either the manifold boundary of $\mathcal{M}$ or the asymptotic/conformal boundary of $\mathcal{M}$). Case (i) is trivial (i.e., $\mathcal{O} = 0$); in case (ii) the observable is completely spread over the entire manifold; and in case (iii) the observable lives at the boundary. (In a way, we may think of the holographic principle as further asserting that case (ii) either does not happen or is equivalent to case (iii).) Accordingly, one of the core structures of QFT is put under stress: there is no microcausality.

From another point of view, it can also be seen that in effective, perturbative quantum gravity, dressed observables at spacelike separated regions (with respect to a background spacetime) always fail to commute [151, 152, 153].

The deeper point is that, in gravity, the notions of subregions and causal structures are fundamentally *relational* and state-dependent. That is, it is clear that we, here at Earth or in our lab, can still make local observations — but they are always in reference to something else, say the corner of the room. But matter is composed of particles, which are quantum excitations, so the references themselves are states in the theory and these "effectively local observables" only exist in small sectors of the full Hilbert space in which the observer/reference "exists". In other words, such an observer may be understood as some class of states displaying some identifiable feature, like a "lump" of coherent field excitations spread over a sufficiently small spatial region (or thin tube in spacetime). It is only within a sector where an observer exists that "approximately local" observables can be defined. The same goes for subregions: the only way to define "where is the boundary" of a subregion is to refer to such an observer. Thus, it is natural to believe that local observables or subregions are concepts that simply do not exist in proper quantum gravity, and are emergent concepts that only make sense in very special small sectors of the full Hilbert space.

But how could one make this idea of observers and relatively-local observables more concrete? What basic properties a state should satisfy in order to contain a observer/reference? That is, what are such “identifiable features”? How to construct dressing with respect to such a feature, given that it is inherently quantum? Or is it that those concepts only make sense in regimes where the observer behaves in a fully classical manner? And how do these concepts manifest at different levels (e.g., full quantum gravity or effective field theory)?

Although it feels unnatural, it is also conceivable that in a full theory of quantum gravity there is some intrinsic prototypical notion of subregions, in the sense that there is an underlying structure to the operators in the theory which organizes them in a way that allows one to reconstruct some notion of topology or causality. In a typical (non-gravitational) quantum field theory, operators are naturally organized as nets of von Neumann algebras (i.e., maps from the directed preordered set of spacetime subregions into the space of von Neumann subalgebras of operators). This is a great amount of additional structure on the space of operators, which would not survive in quantum gravity. But perhaps there could be weaker structures which could survive, e.g., the spectrum of certain important operators such as the total energy (in classes of spacetime with a non-trivial conformal boundary). An interesting work [154] suggests that just the spectrum of the Hamiltonian may contain enough information to reconstruct a notion of locality: generic spectra most likely are not compatible with any local structure, very few spectra are compatible with one local structure, and even fewer are compatible with more than one structure (and those very special cases are then said to exhibit "duality").[78]

We see thus two main routes in trying to understand the role of subsystems. One is what we followed in the work, namely taking the classical notion of subsystems (i.e., causal diamonds[79]) as the basis for quantization, producing of a "piece of quantum spacetime", and then attempting to construct the whole of the quantum spacetime from the pieces; the other is to quantize the whole spacetime and, at the quantum level, try to identify a viable concept of subsystem (even if only emergent, or approximate, in special regimes). The former seems more practical, but the latter seems to me more likely to be correct.

17 On gauge-fixing time

In our reduction process we dealt with "time diffeomorphisms" by a particular gauge-fixing prescription based on CMC slices (and subsequently quotienting out the spatial diffeormorphisms). The reduced phase space is independent of how exactly the reduction process is carried out. In particular, it is independent of the choice of gauge-fixing, provided that the choice is legitimate (i.e., that it is globally well-defined, intersecting every gauge orbit in non-empty, non-disjoint submanifolds). If

[78]The result is derived for finite dimensional Hilbert spaces, and it is unclear whether it generalizes to infinite dimensions. The notion of locality is, roughly speaking, based on whether it is possible to factorize the Hilbert space into a "network of qudits" and express the Hamiltonian as a sum of operators supported in clusters (of limited size) of nearby factors.

[79]There are also other classical notions that could be considered. For example, the region of spacetime that is "operationally accessible" to an observer with a finite lifetime (i.e., the set of events that the observer can send a signal to and receive one back from). This is a subclass of our causal diamonds, consisting of "pointy diamonds", defined as the intersection of the causal past of a point with the causal future of another point (in the past of the former).

there is a consistent canonical quantization, it should therefore not be affected by such a procedure.

However, as mentioned above (Sec. 14), there are some concerns. In particular, canonical quantization is full of ambiguities and there are other features of the phase space, beyond its symplectic structure, that affect what one may consider "natural" choices. For example, the choice of the quantization group may be influenced by how the phase space presents itself. In our case (and similarly any reduction à la Moncrief), the phase space presents itself as a cotangent bundle, $T^*\mathcal{Q}$, where $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, and that structure is additional to the symplectic one. The construction of the canonical group started from a physically motivated symmetry (namely, diffeomorphisms acting non-trivially along the corner of the diamond), but the extension to a transitive group of symplectomorphisms was motivated by the cotangent bundle structure.

Another issue is whether this classical choice of time really breaks the full diffeomorphism invariance of the theory. On the face of it, that might indeed be the case as time remains a real parameter while spatial geometric observables become quantum operators. Philosophically that is a bit disconcerting, for two reasons. First, it seems that physical significance is being assigned to observations made at a sharp, global time. Second, and more importantly, it is unclear how to interpret this time at the quantum level. Classically, given a causal diamond, it is fine to ask "what is the value of a given observable at the slice with constant mean curvature

$\kappa = -\tau$?", but quantum mechanically there is no well-defined spacetime geometry so what does it even mean to ask "what is the result of measuring a given operator at time τ"? One could perhaps make sense of "S-matrix" type of observations, for the limits $\tau \to \pm\infty$ have a simpler (classical) interpretation as corresponding to the past and future horizons, but even the meaning of that is not as clear at the quantum level.

Returning to the point about breaking the diffeomorphism invariance, a fair way to address this question would be to check whether other (intrinsic[80]) gauge-fixings of time also lead to the same quantum theory (or, more precisely, that there is at least one quantum theory in the intersection of all such quantizations). The problem with this approach is that it is exceedingly difficult to find other gauge-fixing prescriptions that are globally well-defined in gravity. In that regard, Moncrief's program may stand at a very unique position. If there are no other gauge-fixings that work properly in gravity, in some useful class of spacetimes, does it endow this mathematical prescription with a physical significance? And what ensures that there are not many other prescriptions which are just not known?

Finally, there is the question of the applicability of the CMC gauge-fixing prescription. In the introduction we pointed out that some spacetimes may not admit CMC slices, and that one possible motivation to consider causal diamonds is that

[80]It seems important that the choice of time is "internal", i.e., defined only from the available (physical) structures within the system. An example of the contrary, leading to failure, is the naive attempt to quantize a single (scalar) relativistic particle in Minkowski, whose action is the proper length of its worldline, $S = \int ds = \int d\lambda \sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}$, by gauge-fixing time using a background time coordinate. In that case, the resulting quantum theory is not compatible with causality as one could define localization (position) operators, in contradiction with Hegerfeldt's theorem [155].

one could apply this rigorous quantization "locally" (in spacetime) and subsequently attempt to construct a larger "quantum spacetime" from these pieces (even if the classical limits of these larger spacetimes do not admit CMCs). One may worry that this separation of scales is not consistent in quantum gravity: even in arbitrarily small regions of spacetime, there could be fluctuations of geometry and topology such that not even local CMC slices exist. It is thus unclear if the quantum spacetime produced by gluing "topologically trivial quantum causal diamonds" is consistent and complete, or whether one needs to explicitly include all "topologically non-trivial quantum causal diamonds" in the sewing kit.

18 On boundary conditions (and gluing)

A potential limitation of our approach to causal diamonds was the choice of the boundary condition, fixing the induced metric on the boundary (or, in higher dimensions, the boundary volume element). This choice was required to justify the applicability of Moncrief-Lichnerowicz program of phase space reduction, specifically so that the CMC gauge was accessible (i.e., generic Cauchy slices could be deformed into CMCs via *gauge flows*).[81] This, of course, attributes a "classical reality" to the boundary length, which is now a fixed real parameter ℓ. While it could still be possible that such a parameter gets "quantized", similarly to how eg was quantized

[81] In (2+1)-dimensions one could avoid this limitation by focusing strictly on boundary charges, which should be sufficient to characterize the phase space since there are no local degrees of freedom. However, such approaches are not generalizable to higher dimensions, while the Moncrief-Lichnerowicz naturally is. In fact, this was our main justification for following this route.

in the case of the particle on the sphere (namely, via the Casimir matching principle), this does not happen for the causal diamond. In fact, ℓ does not even feature explicitly in the $\mathfrak{bms}_3$ algebra. (Nevertheless there are other mechanisms where ℓ could get quantized, as it appears in other physical observables such as the CMC time-evolution Hamiltonian, which is the CMC volume, and certain requirements on the representation of such operators, like lower boundedness of spectrum, could impose restrictions on parameters such as ℓ.)

In any case, it still seems physically inappropriate to fix the boundary metric (or, in higher dimensions, the boundary volume element). For one, we would expect that all aspects of the geometry, including any length or volume elements, to be true quantum observables, that is, described by operators in the theory. Second, to advance on the goal of sewing causal diamonds into a larger spacetime mesh, it seems important to have "free" boundary conditions, so that generic causal diamonds can be attached to each other.

So how could we drop this boundary condition? A natural solution is to employ the idea of *edge modes*, particularly in the context originally described by Donnelly and Freidel [156, 102, 157] (and subsequently by others [52, 158, 159, 160, 161, 162]). Essentially, the idea is to enlarge the (pre-)phase space by introducing "auxiliary" degrees of freedom along the edge, whose function is to modify the symplectic form in such a way as to restore generic gauge transformations into the formalism (even those that are "broken", or rather promoted to non-trivial symmetries, near the boundary). These edge modes can be understood as "minimally-structured observers", or frames of reference, registering the field configurations (and a certain number of normal derivatives) at the boundary. For example, in the case of electromagnetism in a (spatial) region Σ, the "standard" symplectic form is given by

$$\omega_\Sigma = \int_\Sigma \delta E \wedge \delta A$$

where E is the electric flux $(n-1)$-form and A is the magnetic 1-form (considering an $(n+1)$-dimensional spacetime). Naively, general transformations $A \mapsto A + d\lambda$ (and $E \mapsto E$), where $\lambda : \Sigma \to \mathbb{R}$, would be gauge; but it turns out that only those with $\lambda|_{\partial\Sigma} = 0$ are actually gauge.[82] One can then enlarge the phase space by introducing edge modes, $\varphi \in \Omega^0(\Sigma)$ and $\varepsilon \in \Omega^{n-1}(\Sigma)$, adding to the symplectic form a term

$$\omega_{\partial\Sigma} = \int_{\partial\Sigma} \delta\varepsilon \wedge \delta\varphi$$

where φ and ε can be respectively interpreted as "measuring" the values of λ and the normal electric flux, $\vec{E} \cdot \vec{n}$, at the boundary. The enlarged phase space, $\widehat{\mathcal{P}}_\Sigma$, is defined as a "fusion product",

$$\widehat{\mathcal{P}}_\Sigma = \mathcal{P}_\Sigma \times_{H_\partial} \mathcal{P}_{\partial\Sigma}$$

[82] This is assuming that there is an (omitted) coupling to charged matter. In particular, if λ approaches a non-zero constant at the boundary, the Hamiltonian charge associated with this flow is simply (proportional to) the electric charge contained in the region (assuming $\delta\lambda = 0$). If there is no matter, then any λ that approaches a constant at the boundary is also gauge, but those that approach a non-constant at the boundary are still not gauge.

where $H_\partial := \varepsilon - E|_{\partial\Sigma} \approx 0$ imposes the "flux constraint", and the total symplectic form is

$$\widehat{\omega}_\Sigma = \omega_\Sigma + \omega_{\partial\Sigma}$$

It can be verified that the flow in $\widehat{\mathcal{P}}_\Sigma$ defined by $A \mapsto A + d\lambda$ and $\varphi \mapsto \varphi - \lambda$ (and $E \to E$ and $\varepsilon \mapsto \varepsilon$) is indeed gauge. The result of this construction is to "restore" the naive gauge transformations, while still preserving the original symmetries. In particular, the symmetries formerly associated with λ's that approach non-zero values at the boundary now correspond to a flow that changes only the edge modes, $A \to A$ and $\varphi \to \varphi + \alpha$ (and $E \to E$ and $\varepsilon \mapsto \varepsilon$), where $\alpha : \partial\Sigma \to \mathbb{R}$. Note that, in this form, it is quite explicit that the symmetries are fundamentally associated with boundary fields.

The story is very similar in gravity, with the diffeomorphism gauge being "broken" at the boundary, but restored with the introduction of appropriate edge modes. In this case, we may think of the edge modes as providing the minimal structure to delineate the location of the corner of the diamonds, as well as its "orientation" (i.e., a orthonormal frame on it). Not all these edge modes are necessary for our purposes: we only need the *corner boosts* in order to reintroduce the gauge directions that allow generic deformations of Cauchy slices. Similarly to the case of electromagnetism, the "edge mode phase space" (associated with boosts), $\mathcal{P}_{\partial\Sigma}$, consists of a *volume element* $(n-1)$-form ε on $\partial\Sigma$ (in an $(n+1)$-dimensional spacetime), analogous to the electric flux, and a conjugate scalar *hyperbolic angle* variable φ (registering the "tilt" of the Cauchy slice as it intersects with the corner), analogous to the boundary $U(1)$ EM phase. The "flux constraint", in the fusion product, here just enforces that ε is induced from the bulk geometry, that is, $H_\partial := \varepsilon - i_n \vartheta_h|_{\partial\Sigma} \approx 0$, where ϑ_h is the spatial volume element (associated with the metric h) and n is the unit vector normal to the boundary (tangent to the Cauchy slice). With that, one could carry out the Moncrief-Lichnerowicz reduction in the "bulk phase space", $\mathcal{P}_\Sigma$, and have the remaining edge modes neatly accounted for in $\mathcal{P}_{\partial\Sigma}$. Roughly speaking, the process would thus be to first turn $\mathcal{P}_\Sigma$ into $\mathcal{P}_\Sigma \times_{H_\partial} \mathcal{P}_{\partial\Sigma}$, and then reduce it to $\widetilde{\mathcal{P}_\Sigma} \times_{H_\partial} \mathcal{P}_{\partial\Sigma}$.

Another benefit of this construction is, as alluded before, the natural manner in which different subregions can be glued together. The point is that, in non-abelian gauge theories, the set of observables in a union of regions is larger than what can be constructed from observables in each region separately. In particular, a Wilson line passing through (and only through) regions Σ and Σ' is supported in $\Sigma \cup \Sigma'$, but cannot be constructed from observables in Σ and Σ' alone. At the classical level, the phase space of $\Sigma \cup \Sigma'$ can be defined by another fusion product between the enlarged phase spaces of each region. Namely, if Σ and Σ' intersect only at their boundaries,

$$\widehat{\mathcal{P}}_{\Sigma\cup\Sigma'} = \widehat{\mathcal{P}}_\Sigma \times_{G_S} \widehat{\mathcal{P}}_{\Sigma'} = (\mathcal{P}_\Sigma \times_{H_\partial} \mathcal{P}_{\partial\Sigma}) \bowtie_S \left(\mathcal{P}_{\Sigma'} \times_{H'_\partial} \mathcal{P}_{\partial\Sigma'}\right)$$

where $S = \partial\Sigma \cap \partial\Sigma'$ and $\bowtie_S$ denotes the fusion product where the "fluxes" are matched on S, $H_S := (H_\partial - H'_\partial)\,|_S \approx 0$, and the flow generated by H_S is quotiented over. A very similar "fusion product" construction also applies to the quantum level.

19 On diffeomorphism charges

The Hamiltonian charges associated with diffeormorphisms that act non-trivially near the boundary of spacetime play a key role in many discussions of quantum gravity. A notable example is the work of Brown and Henneaux [26] showing that there is a large group of symmetries associated with "large diffeomorphisms" in asymptotic AdS_3 and, when realized as Hamiltonian charges on the phase space, the algebra acquires a non-trivial central charge $c = 3\ell/2G$. This result is regarded as one of the first hints towards the holografic principle for the extended group, $\mathit{Vira} \times \mathit{Vira}$, is also the group of symmetries of a 2-dimensional CFT. Furthermore, it was later discovered [163] that, by assuming modular invariance and applying Cardy's formula, one correctly obtains the Bekenstein-Hawking entropy for the BTZ black role. This group was also used in [30] as the basis for the reduced phase space construction: as there are no bulk degrees of freedom, the "large diffeomorphisms" should act transitively on the reduced phase space (i.e., the degrees of freedom are only "edge modes", ignoring topological modes that could exist if the Cauchy slice were non-trivial).

It is thus natural to consider the role of boundary diffeomorphisms in causal diamonds. In fact, as we discussed above (Sec. 18), they are central to the formalism of edge modes [156, 102, 157, 52, 158, 159, 160, 161, 162]. Very recently, the boundary diffeomorphism charges have been computed, from a covariant phase space perspective, for causal diamonds in $(2+1)$-dimensions [164]. These charges have nice and simple geometric interpretations.

It should be pointed out, however, that there is an important difference between asymptotic AdS_3 (or other asymptotic spacetimes) and causal diamonds. In the former case, the group of asymptotic diffeomorphisms does act legitimately on the reduced phase space. In fact, the asymptotic diffeomorphisms are defined precisely as the set of diffeomorphisms that preserve the class of asymptotic spacetimes under consideration, and thus they always map a state into another state. In the case of causal diamonds this does not happen, at least when normal translations of the corner are allowed. In fact, as these boundary diffeomorphisms act by deforming the diamond corner, there is not preventing them from "pushing it too far", that is, deforming a "nice corner" into a loop that is not acausal or that develops crossings and knots, and thus can no longer be the corner of another diamond. In other words, the Hamiltonian flow of generic boundary diffeomorphism charges will not be complete, reaching the "end of phase space" in finite parameter-length. While these charges do cover the phase space with regular coordinates, they are not appropriate "canonical charges" to be used as the basis of quantization, at least from Isham's perspective. Equivalently, it would seem that the unitary representations of the corner group are not (entirely) relevant for the "quantum theory" of this system.

The situation is analogous to quantizing a particle on the half-line, as we explained in Sec. 15.2: the coordinates $x > 0$ and p cover the phase space and also form a closed algebra (together with the constant function 1), but they do not exponentiate to a group acting *on* the phase space and thus they do not lead to a sensible quantum theory (namely, the spectrum of x is not contained in $\mathbb{R}^+$); instead, one could use the affine group to quantize the system, whose charges are x and xp.

We may have a possible "trade-off" at hand in the case of causal diamonds: while boundary diffeomorphism charges are physically natural and have a nice geometric interpretation, then do not form a viable basis for quantization; on the other hand, the $\mathfrak{bms}_3$ charges that we have constructed were derived from a group of symmetries that acts properly on our phase space, and thus can be used for quantization, but unfavorably half of them (namely, the Q charges) seem to lack a simple geometric interpretation.

20 On symplectic ambiguities

As discussed before (Sec. 14), the symplectic structure is not classically observable, as it depends on a choice of action. Usually, the action is sufficient for uniquely determining the symplectic form, but the situation can be subtle in field theories defined in certain spacetimes. (As we will emphasize later, the issue is not with field theories per se, but with dynamically open systems.) In particular, there might be ambiguities affecting boundary terms in the symplectic form, which can most readily be seen from the covariant perspective. The action in an $(n+1)$-dimensional spacetime (region) $\mathcal{M}$ is given by

$$S = \int_{\mathcal{M}} L + \int_{\partial\mathcal{M}} l$$

where L is the bulk spacetime $(n+1)$-form Lagrangian and l is the boundary spacetime n-form lagrangian. Generally, $\delta L = E\delta\phi + d\Theta$, where $E \approx 0$ imposes the equations of motion and Θ is a spacetime n-form and field-space 1-form. It is often taken [165] that the symplectic form is obtained from Θ as

$$\omega = \int_{\Sigma} \delta\Theta$$

where Σ is a Cauchy slice. However, as can be immediately noticed, Θ is only defined (from δS) up to the addition of a closed spacetime n-form. For simplicity, let us assume trivial spacetime topology so that this closed form is necessarily exact, i.e., equal to $d\alpha$ for some spacetime $(n-1)$-form α. This alternative choice, $\Theta' = \Theta + d\alpha$, consequently defines a different symplectic form

$$\omega' = \int_{\Sigma} \delta\Theta' = \int_{\Sigma} \delta\Theta + \int_{\Sigma} d\delta\alpha = \omega + \int_{\partial\Sigma} \delta\alpha$$

where it was used that $d\delta = \delta d$. This reveals an ambiguity in the symplectic structure, particularly affecting the "boundary component" of the symplectic form.

In certain scenarios, one can use well-motivated physical principles to resolve this ambiguity. A nice situation is described by Harlow and Wu [55], considering "cylinder-shaped", *dynamically-closed* regions of spacetime. The principle invoked is simply that the "stationary-action principle is well-posed". That is, if $\partial\mathcal{M}$ is decomposed as a past Cauchy slice Σ_-, a (timelike) cylinder boundary Γ and a future Cauchy slice Σ_+, and one considers a class of field configurations ϕ where $\phi|_{\Sigma_\pm} =: \phi_\pm$ are fixed (as standard in setting up the variational problem in mechanics) and ϕ (and possibly some number of derivatives) satisfy all the a priori boundary conditions on Γ, then the variational principle is well-posed if $S[\phi]$ has at least one stationary configuration. In general, the variation of S gives

$$\delta S = \int_{\mathcal{M}} E\delta\phi + \int_{\Sigma_+} (\Theta + \delta l) - \int_{\Sigma_-} (\Theta + \delta l) + \int_{\Gamma} (\Theta + \delta l)$$

where the orientation of $\Sigma_\pm$ are chosen to be the same. The first condition for a stationary point is that $E = 0$, so the equations of motion must be solvable. For a generic class of theories and boundary conditions, the integral on Γ could prevent a stationary configuration from existing.[83] The situation for the integrals over $\Sigma_\pm$ is generally much better since the imposition that $\delta\phi|_{\Sigma_\pm} = 0$ is often sufficient for ensuring that they vanish. It is thus reasonable to assume that, in a well-posed theory, things must work out such that $\int_\Gamma (\Theta + \delta l)$ contributes, at most, with

[83]For example, in a free scalar field theory, $L = \frac{1}{2} d\phi \wedge \star d\phi$ (and $l = 0$), one has $\Theta = \star d\phi\, \delta\phi$. Unless the boundary conditions are Dirichlet or Neumann, $\Theta|_\Gamma$ is non-trivial and, by infinitesimally deforming ϕ (around the kernel of E) with a tiny but sharp bump as it approaches Γ, it is possible to make δS arbitrarily large.

boundary terms (since $\partial\Gamma = \partial\Sigma_+ \cup \partial\Sigma_- \subset \Sigma_+ \cup \Sigma_-$). A simple and quite generic way to have this is if

$$(\Theta + \delta l)\,|_\Gamma = dC$$

for some spacetime $(n-1)$-form and field-space 1-form C. The view here is that the "inputs" of the physical theory consist of specifying L (which is associated with the bulk dynamics) and the boundary conditions, and then an l and a C are, if possible, introduced a posteriori with the purpose of making the variational principle well-posed. In this case, note that the on-shell first-order variation of S is

$$\delta S \approx \int_{\Sigma_+} (\Theta + \delta l) - \int_{\partial\Sigma_+} C - \int_{\Sigma_-} (\Theta + \delta l) + \int_{\partial\Sigma_-} C$$

From a logic similar to what was discussed in Sec. 7.1, it is justified to identify the symplectic potential as the terms at the "future time" (i.e., $\Sigma_+ \cup \partial\Sigma_+$), from which the symplectic form is simply the (field-space) exterior derivative,

$$\omega = \delta\left(\int_{\Sigma_+} (\Theta + \delta l) - \int_{\partial\Sigma_+} C\right) = \int_{\Sigma_+} \delta\Theta - \int_{\partial\Sigma_+} \delta C$$

As the condition $(\Theta + \delta l)\,|_\Gamma = dC$ is generally "hard" to satisfy, the ambiguity in Θ being replaced by $\Theta + d\alpha$ would force a unique different choice of C, namely $C + \alpha$. That is, the symplectic form ω is, in this case, insensitive to the ambiguity.

A similarly well-motivated and sufficiently constraining prescription is, however, not fully developed for causal diamonds. A fundamental difference may be in that causal diamonds are inherently not dynamically closed: things can go in (through the past horizon) or out (through the future horizon). There has been some effort in proposing resolutions for these ambiguities in open systems, usually assuming some sort "symplectic flux condition" through the horizons [56, 53]. It is unclear to me how natural they are (or how can they be chosen) in the context of causal diamonds. An interesting but quite different approach is proposed in [57], where they adopt an "a priori path integral" perspective. My intuition is, however, that the path integral is simply a particular realization of the representation theory of canonical observables, and can only be justifiable a posteriori, at least in the context of canonical quantization (in particular, I am not questioning the consistency or validity of a point of view in which path integrals are truly axiomatic).

In any case, the point is that the reduced phase space for causal diamonds, and particularly the structure of edge modes, is subject to this ambiguity in the symplectic form. In the absence of a good prescription for resolving the ambiguity, it appears that the only reasonable approach is to generalize the formalism to include all possible choices. It may be that the correct choice of symplectic structure will only become clear after all causal diamonds are glued together into a dynamically-closed universe, say to guarantee the consistency of the whole theory or reproduce physically expected results (e.g., the proper counting for black hole microstates).

21 On matter

Another interesting generalization would be to include matter. For simplicity, one could start with simple additions like a (minimally-coupled) scalar field or electromagnetism. With matter one has access to more interesting observables and, especially in (2+1)-dimensions, a more realistic physical system containing "quasi-local" observers, as described in Sec. 16. (In higher dimensions one could, in principle, use "lumps" of gravitational disturbances as observers.[84])

Technically speaking, adding matter into our formalism is viable as long as the main premises of the Moncrief-Lichnerowicz program are not nullified. In particular, it must be that (i) for generic classical states the causal diamond still admits a CMC foliation, (ii) the components of the stress-energy tensor scale in a definite way under conformal transformations, so that a Lichnerowicz equation can be used to solve the Hamiltonian constraint ($\mathcal{G}_{00} + \Lambda g_{00} = 8\pi \mathcal{T}_{00}$), and ($iii$) this Lichnerowicz equation possesses the desired existence and uniqueness properties for the solutions.

On point (i), looking at the arguments in Sec. 4.1, we see that the main assumption on the (local) geometry was the *timelike convergence condition*, i.e., $\mathcal{R}_{ab} u^a u^b \geq 0$ for all timelike vectors u. This is essentially the assumption that test particles are "attracted" to each other (i.e., congruences of geodesics are focusing). To ensure this, for all classical states, the matter content of the theory should

[84] While there are no quasi-local observers in $(2+1)$-dimensions, the "pure gravity" quantum theory should still be a good learning ground for the meaning of geometry, as the classical theory still contains the basic notion of "spacetime shape", and to understand the fate of that concept in the quantum theory is valuable.

satisfy the *strong energy condition*. That is,

$$\left(\mathcal{T}_{ab} - \frac{1}{n-1}\mathcal{T} g_{ab}\right) u^a u^b \geq 0$$

for all timelike vectors u. (Again, the dimension of the spacetime is $n+1 \geq 3$.) For example, a free massless scalar field and a free electromagnetic field satisfy this condition.

On point (ii), there needs to be a choice of conformal rescalings of the components of the matter field which preserves all the constraints, except possibly the Hamiltonian constraint. Recall that the Lichnerowicz method is mostly useful as a tool for solving the Hamiltonian constraint: it is assumed that all the other constraints are "solved first" and that their solutions are not "destroyed" when conformally deforming the fields in order to solve the Hamiltonian constraint (see Sec. 3.2). In the case of electromagnetism, we wish to preserve the $U(1)$ constraint

$$\nabla_i F^{0i} = \nabla_i E^i = 0$$

and the momentum constraint

$$-2\nabla_i \pi^{ij} = 16\pi\sqrt{h}\mathcal{T}^{0j}$$

where the stress-energy tensor is

$$\mathcal{T}_{ab} = \frac{1}{4\pi}\left(F_{ac}F_b{}^c - \frac{1}{4}g_{ab}F_{cd}F^{cd}\right)$$

In $3+1$ spacetime dimensions, this case was analyzed in [67], revealing that electromagnetism does admit appropriate conformal scalings. The corresponding transformations on the components of the stress-energy tensor, under a conformal transformation $h_{ij} \mapsto \phi^4 h_{ij}$, end up being $\mathcal{T}^{0j} \mapsto \phi^{-4}\mathcal{T}^{0j}$ and $\mathcal{T}_{00} \mapsto \phi^{-8}\mathcal{T}_{00}$. The associated Lichnerowicz equation is

$$\nabla^2\phi - R\phi + \sigma_{ij}\sigma^{ij}\phi^{-7} - \frac{3}{8}\tau^2\phi^5 + 16\pi\mathcal{T}_{00}\phi^{-3} = 0$$

where $\Lambda = 0$ (but including a $\Lambda \neq 0$ would be trivial). The generalization to other dimensions is also straightforward.

On point (*iii*), one needs to ensure that the associated Lichnerowicz equation has one and only solution for each seed data (which now includes also matter fields). For the case above, [67] discusses conditions for existence and uniqueness. Electromagnetism turns out to be very agreeable with the Lichnerowicz method, as it also satisfies the *weak energy condition* (i.e., $\mathcal{T}_{ab}u^a u^b \geq 0$ for all timelike vectors u), thus at least marginally satisfying the conditions described in the paper. (By "marginally" I mean that it only potentially fails in very special circumstances, say when $\sigma_{ij}\sigma^{ij} + \mathcal{T}_{00}$ is everywhere zero or $\tau = 0$.) The further inclusion of a non-positive cosmological constant should only make the situation better, possibly ensuring that conditions for existence and uniqueness are strictly satisfied.

In $(2+1)$-dimensions, there is also another kind of matter: conical singularities. These are interesting because they are purely "geometrical particles", described by the standard Einstein-Hilbert action except for an enlargement on the space of metrics (relaxing the usual requirement on smoothness by allowing for these pointwise singularities). A *conical singularity* is a type of metric singularity that locally resembles a "cone tip".[85] That is, a 2-dimensional manifold contains a conical singularity at the point c if: (i) there is a neighborhood U of c such that the metric in $U-\{c\}$ is regular, and (ii) given the family of metric balls $B(r;c)\subset U$ of (geodesic) radius r centered at c, covered by polar coordinates (r,θ), with $\theta\in(0,2\pi]$, then the metric restricted to $B(r;c)$ approaches $dr^2+\kappa^2r^2d\theta^2$, as $r\to 0$, for some constant $\kappa>0$. A natural way to characterize the "value" of the singularity is to specify how much the perimeter of a small circle of radius r, centered at c, differs from $2\pi r$. Thus we define the *angle anomaly* by

$$\Delta := \lim_{r\to 0}\frac{length[\partial B(r;c)]}{r}-2\pi$$

In terms of κ, we have $\Delta = 2\pi(\kappa-1)$. A "familiar cone", embedded in the 3-dimensional Euclidean space, with aperture angle α, will have $\kappa=\sin\alpha$ and thus angle anomaly $\Delta = 2\pi(\sin\alpha-1)<0$, that is, it is an *angle deficit*.

[85] Alternatively, instead of thinking of "singular metrics", it is also possible to consider regular metrics on topological punctured discs.

In gravity, angle deficits are physically more "natural" as they correspond to particles of positive mass. They have been extensively studied, especially in the context of AdS_3 [166, 167, 168, 169]. In particular, for pure gravity, the (classical) energy spectrum of the theory consists of the vacuum AdS_3 solution with energy $-1/8G$, BTZ black holes [170] with positive mass $(r_+^2 + r_-^2)/8G\ell^2$ and, filling this gap, conical singularities with angle deficits; the conical singularities with angle surpluses would sit below the vacuum (consistent with the interpretation that they have negative mass). It would be interesting to investigate the consequences of their inclusion in our causal diamonds. In particular, they would act as simple observers, introducing a larger class of observables. For example, one could ask what is its "lifetime" (i.e., the time it takes to go from the past to the future horizon) or its relative position with respect to the corner of the diamond.

In our approach, conical singularities could (presumably) be implemented in two ways. Following the conformal coordinates approach of Sec. 6, the first way is to replace the reference disc with one containing a conical singularity at the center,

$$\bar{h} = dr^2 + \kappa^2 r^2 d\theta^2$$

and consider smooth conformal transformations to the physical disc. While this approach can be implemented in a straightforward manner, the disadvantage is that the number and masses of the conical singularities are fixed as classical inputs. The second way would be to keep the reference disc as the regular unit round disc, but allow for conformal transformations that diverge at isolated points. With such divergences it is possible to "create" a conical singularity. For example, consider a Weyl transformation $dr^2 + r^2\theta^2 \mapsto \Omega(r)^2(dr^2 + r^2\theta^2)$. With respect to the deformed metric, the radius of a circle at coordinate r is $\int_0^r dr'\Omega(r')$ and its perimeter is $2\pi r\Omega(r)$. The angle anomaly is thus

$$\Delta = \lim_{r\to 0} \frac{2\pi r\Omega(r)}{\int_0^r dr'\Omega(r')} - 2\pi = 2\pi \left.\frac{d\ln\Omega}{d\ln r}\right|_{r=0}$$

where L'Hôpital's rule was used. It is clear that if Ω is continuous (and nowhere zero) then no conical singularity will be introduced, but if, say, $\Omega(r) = r^{\kappa-1}$, then an angle anomaly $\Delta = 2\pi(\kappa - 1)$ will be introduced at the origin. Note that for conical deficits Ω is nowhere zero but diverges to $+\infty$ at the origin, while for conical surpluses Ω is bounded but goes to zero at the origin. In this way, it seems more natural to address the case where the conical singularities are dynamically generated, with variable numbers and masses. It is curious to point out that, in the context of asymptotic AdS_3, Maloney and Witten [30] noticed that a theory of pure gravity (without conical singularities) is somewhat "sick" (in particular, with a physically unsound modular-invariant partition function), but then others (including Maloney) [171] realized that the inclusion of some conical singularities (with a quantized angle deficit $\Delta = 2\pi(1/n - 1)$, where $n \in \mathbb{N}$) offers a potential "cure" for this issue. It would be interesting to see whether, just in the context of causal diamonds, one can also encounter a similar quantization of the mass spectrum of conical singularities.

Appendices

A Glossary, symbols and conventions

This appendix features a quick reference guide to recurrent terms and symbols found in this paper. It is organized roughly in the order of appearance in the text, but also in such a way that later entries only refers to terms and symbols already introduced in earlier entries. Each entry is indicated by an *underlined italic name* followed by the explanation; in cases where relevant symbols are introduced in the explanation, those symbols are displayed on the left margin for easier reference. Despite our efforts to keep the notation uniform throughout, there may be instances where a symbol is used with a different meaning for a specific section. For example, Λ is mostly used to denote the cosmological constant, but sometimes it can be used to denote a Weyl scaling factor. There are also some variations between Part I and II. For example, Ψ is used in Part I to denote diffeormorphisms of the disc, while in Part II it is used to denote quantum states (particularly in their wavefunction realization). We hope that the context will prevent confusion in those instances. Some general conventions are also referenced in this section.

Units: We adopt units in which the speed of light and Planck's constant are 1,

$$c = \hbar = 1$$

Signature: The spacetime metric is assumed to be of signature

$$- + + \cdots$$

<u>*Causal diamonds*</u> The domain of dependence of an acausal spacelike disc is referred to as a causal diamond.

$T^a_{\ b}$, $T^\mu_{\ \nu}$, $T^i_{\ j}$ <u>*Tensors*</u> The abstract index notation is used for tensors, where subscript Latin letters indicate covariant tensor slots and superscript Latin letters indicate contravariant slots (typically we reserve for this purpose letters from the first half of the alphabet like a, b, c etc).[86] A basis of vectors is typically denoted as e_μ, or $(e_\mu)^a$, where the subscript label μ runs from 1 to the dimension d of the manifold; the associated dual basis is denoted by e^μ, or $(e^\mu)_a$, with a superscript label, and defined so that $e^\nu(e_\mu) = \delta^\nu_\mu$, where δ^ν_μ is the Kronecker delta; tensors can be decomposed in that basis with components $T^{\nu_1\nu_2\cdots\nu_n}_{\mu_1\mu_2\cdots\mu_m} := T(e_{\mu_1}, e_{\mu_2}, \cdots e_{\mu_m}; e^{\nu_1}, e^{\nu_2}, \cdots e^{\nu_n})$. Typically we reserve Greek letters (μ, ν, etc) to denote tensors decomposed in a spacetime basis, and Latin letters from the second half of the alphabet (like i, j, etc) for tensors decom-

[86] If V denotes the space of vectors at a point p on a manifold, and V^* denotes the dual vector space (i.e., space of 1-forms, or linear maps from V to $\mathbb{R}$) at p, then a tensor T at p of type $({}^n{}_m)$ is a linear map from $V^m \times (V^*)^n$ to $\mathbb{R}$ and it is denoted by $T^{b_1b_2\cdots b_n}_{a_1a_2\cdots a_m}$; a 1-form α is a linear map from V to $\mathbb{R}$ and therefore has one covariant index and is denoted as α_a; a vector ξ can be seen as a tensor with one contravariant index, denoted by ξ^a, because of the natural duality between vectors and double dual vectors ($V \sim V^{**}$) which identifies ξ with a linear map from V^* to $\mathbb{R}$ (i.e., $\alpha \mapsto \alpha(\xi)$). In this notation, the tensor T applied to m vectors ξ_1, ξ_2, $\cdots \xi_m$ and n dual vectors α^1, α^2, $\cdots \alpha^n$ is expressed by contracting the respective indices, $T(\xi_1, \xi_2, \cdots \xi_m; \alpha^1, \alpha^2, \cdots \alpha^n) = T^{b_1b_2\cdots b_n}_{a_1a_2\cdots a_m}(\xi_1)^{a_1}(\xi_2)^{a_2}\cdots(\xi_m)^{a_m}(\alpha^1)_{b_1}(\alpha^2)_{b_2}\cdots(\alpha^n)_{b_n}$.

posed on a spatial basis. In Part II, most tensors are associated with the phase space, configuration space or groups and algebras, and these will typically be denoted without any indices.

f_*, f^* *Push-forward and pull-back* Given a smooth map between two manifolds, $f : \mathcal{M} \to \mathcal{N}$, the push-forward operator mapping vectors (or contravariant tensors) on $\mathcal{M}$ to $\mathcal{N}$ is denoted by f_* and the pull-back operator mapping 1-forms (or covariant tensors) from $\mathcal{N}$ to $\mathcal{M}$ is denoted by f^*. If f is a diffeomorphism (smooth invertible map) then $f^* = f_*^{-1}$ and any tensor can be pushed forward or pulled back.

Λ *Cosmological Constant* The cosmological constant will be denoted by Λ. In this paper we assume $\Lambda \leq 0$. (In some instances the symbol Λ may be used for other purposes, such as denoting certain maps or Weyl factors.)

ℓ, ℓ_{AdS}, ℓ_P *Length scales* The length of the diamond corner is ℓ, the Anti-de Sitter radius is $\ell_{AdS} = 1/\sqrt{-\Lambda}$ (if $\Lambda < 0$), and the Planck length is $\ell_P = \hbar G$.

Σ, D *Cauchy slice* A Cauchy slice will typically be denoted by Σ. In this paper it is assumed to have the topology of a two-dimensional disc, D. (Because of this, sometimes D is also used to denote the Cauchy slice.)

$\partial\Sigma$, ∂ *Corner* The boundary of any Cauchy slice, $\partial\Sigma$, is referred to as the corner of the diamond. It has the topology of a circle, S^1. (Sometimes we shall abbreviate and simply denote it by ∂ — not to be confused with partial differentiation.)

h_{ab}, g_{ab} — *(Physical) Metric* Metrics in this paper typically refer to the spatial disc Σ and are denoted by h_{ab}. Spacetime metrics are normally denoted by g_{ab}.

$\bar{h}_{ab}$ — *(Reference) Metric* The metric on the reference round unit disc, $dr^2 + r^2 d\theta^2$, is denoted by $\bar{h}_{ab}$. (The bar notation will often be used to indicate objects associated with the reference disc, like $\bar{\sigma}^{ab}$ below.)

γ_{ab} — *Boundary metric* We consider "Dirichlet condition" for the metric induced on the corner, $h_{ab}|_{\partial\Sigma} = \gamma_{ab}$.

$\mathrm{Riem}(\Sigma, \gamma)$ — *Space of metrics* The space of all (sufficiently regular) Riemannian metrics on a manifold Σ is denoted by $\mathrm{Riem}(\Sigma)$. Its subspace consisting of metrics satisfying the boundary condition $h|_\partial = \gamma$ is denoted by $\mathrm{Riem}(\Sigma, \gamma)$.

ℓ, ℓ_{AdS}, ℓ_P — *Length scales* There are three main length scales: the corner length ℓ (determined by γ_{ab}), the Anti-de Sitter radius $\ell_{AdS} = 1/\sqrt{-\Lambda}$ (if $\Lambda < 0$), and the Planck length $\ell_P = \hbar G$. (The Planck constant is introduced here even though it is only relevant for the quantum part, as we shall consider units where $c = \hbar = 1$.)

∇, $\bar{\nabla}$, $\boldsymbol{\nabla}$ — *Covariant derivatives* The covariant derivative on a Cauchy slice associated with a spatial metric h is denoted by ∇; and the one associated with the reference metric $\bar{h}$ by $\bar{\nabla}$. The covariant derivative on spacetime associated with a metric g is denoted by $\boldsymbol{\nabla}$.

$\pounds$, δ, $\imath$ *Lie, exterior and interior derivatives* The Lie derivative along a vector field X is denoted by $\pounds_X$. The exterior derivative of a form is denoted by δ; and the interior derivative (a.k.a., interior product), with respect to a vector X, by $\imath_X$.

(Ψ, Ω), $\mathrm{Con}(\Sigma)$ *Conformal transformation* A transformation on metrics labeled by a pair (Ψ, Ω), where Ψ is a diffeomorphism and Ω is a positive scalar, that acts by multiplying the metric by the scalar (a.k.a., Weyl factor) and pushing-forward by the diffeomorphism, $(\Psi, \Omega)h_{ab} := \Psi_* \Omega h_{ab}$ is called a conformal transformation. Since $\Omega > 0$, we often write it as e^{λ} or e^{ϕ}, where λ, $\phi \in \mathbb{R}$. The space of conformal transformations acting on metrics on a manifold Σ is denoted by $\mathrm{Con}(\Sigma)$. (In the literature the term "conformal transformation" is sometimes used in a restricted sense equivalent to our definition of a "conformal isometry" — see below.)

$\mathrm{Con}(\Sigma^*)$ *Boundary-trivial conformal transformation (BTCT)* A conformal transformation whose multiplicative scalar is equal to 1 at the boundary of the manifold and diffeomorphism acts as the identity at the boundary is said to be a boundary-trivial conformal transformation or BTCT. The space of BTCTs acting on metrics on a manifold Σ is denoted by $\mathrm{Con}(\Sigma^*)$.

Conformal equivalence Two metrics related by a boundary-trivial conformal transformation will be said to be conformally equivalent.

$\mathrm{ConGeo}(\Sigma)$ *Conformal geometries* The space of conformally-equivalent (see above) metrics on a manifold Σ, denoted by ConGeo(Σ), is called the space of conformal geometries on Σ.

ConIso(h) *Conformal isometry* A conformal transformation that leaves a given metric h_{ab} invariant, $h_{ab} = \Psi_* \Omega h_{ab}$, is said to be a conformal isometry. The space of conformal isometries of a metric h is denoted by ConIso(h).

K^{ab}, σ^{ab}, τ *(Physical) Extrinsic curvature* The extrinsic curvature of the Cauchy slice (as embedded in the spacetime) is normally denoted by K^{ab}, its trace part by $-\tau := K := K^{ab}h_{ab}$, and its traceless part by $\sigma^{ab} := K^{ab} - \frac{1}{2}Kh^{ab}$.

$\bar{\sigma}^{ab}$ *(Transformed) Extrinsic curvature* The "conformal pull-back" of the (traceless part of the) extrinsic curvatures to the reference disc are typically denoted by $\bar{\sigma}^{ab}$. If (Ψ, Ω) maps the metric on the reference disc to the physical metric, the $\bar{\sigma}^{ab} = \Omega^2 \Psi^* \sigma^{ab}$. It is usually understood that the momentum constraint is satisfied, so $\bar{\sigma}^{ab}$ are traceless and divergenceless (with respect to $\bar{h}$) symmetric tensors.

χ, $\bar{\chi}$ *"Chi parameter"* A parameter featuring in the Lichnerowicz equation, that appears often in the text, is defined as $\chi = -2\Lambda + \tau^2/2$. In Sec. 8 a dimensionless version of the parameter is introduced, $\bar{\chi} = (\ell/2\pi)^2 \chi$.

ψ, *Diff*$^+(S^1)$ *Boundary diffeomorphisms and its group* The restriction to the boundary circle of a diffeomorphism on the disc is normally denoted by the lowercase version of the symbol, $\psi = \Psi|_\partial$. Similarly, an extension to the disc of a boundary diffeomorphism is denoted by the uppercase version of the symbol. The group of (orientation-preserving) diffeomorphisms on the circle is denoted as $\mathit{Diff}^+(S^1)$. The identity element is denoted by I.

$\mathfrak{diff}(S^1)$, $\mathfrak{diff}^*(S^1)$ *Algebra of boundary diffeomorphisms* The Lie algebra of $\mathit{Diff}^+(S^1)$ is denoted as $\mathfrak{diff}(S^1)$ and consists of vector fields ξ on S^1. The dual Lie algebra of $\mathit{Diff}^+(S^1)$ is denoted by $\mathfrak{diff}^*(S^1)$ and consists of quadratic forms α on S^1. The pairing is $\alpha(\xi) := \int \alpha(\theta)d\theta^2(\xi(\theta)\partial_\theta) = \int d\theta\alpha(\theta)\xi(\theta)$.

$PSL(2,\mathbb{R})$, $\mathfrak{psl}(2,\mathbb{R})$ *Projective special linear group and algebra* The subgroup of $\mathit{Diff}^+(S^1)$ corresponding to the restriction of the (diffeomorphism part of the) group of conformal isometries of the unit round disc to the boundary defines $PSL(2,\mathbb{R})$. Its Lie algebra, denoted $\mathfrak{psl}(2,\mathbb{R})$, consists of the span of vectors ∂_θ, $\sin\theta\,\partial_\theta$, $\cos\theta\,\partial_\theta$

$\mathring{\sigma}$, $\mathring{\mathfrak{diff}}^*(S^1)$ *"Sigma-circle"* The subspace of $\mathfrak{diff}^*(S^1)$ that annihilates $\mathfrak{psl}(2,\mathbb{R})$ is denoted by $\mathring{\mathfrak{diff}}^*(S^1)$ and its elements are typically denoted by $\mathring{\sigma}$. There is a natural correspondence between $\mathring{\sigma}$'s and $\bar{\sigma}^{ab}$ that satisfies the constrains of general relativity. The notation is supposed to bring to mind that "sigma-circle" describes a (traceless) extrinsic curvature "sigma-bar" as a quantity living on the boundary "circle", S^1.

$\mathcal{P}$, $\widehat{\mathcal{S}}$, $\widetilde{\mathcal{P}}$ *Phase space* The pre-phase space (described by spatial metrics and extrinsic curvatures) for the causal diamonds is denoted by $\mathcal{P}$, and the fully reduced phase space by $\widetilde{\mathcal{P}} = T^*[\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})]$. In some occasions it is convenient to work with the partially reduced phase space $\widehat{\mathcal{S}} = \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$.

J, q $\underline{\mathit{PSL}(2,\mathbb{R})\ \mathit{projections}}$ The quotient by $PSL(2,\mathbb{R})$ from $\widehat{\mathcal{S}}$ to the reduced phase space $\widetilde{\mathcal{P}}$ is denoted by J. The quotient by $PSL(2,\mathbb{R})$ from $\mathit{Diff}^+(S^1)$ to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ is denoted by q.

Ω, $\widetilde{\omega}$ *Symplectic form* The pre-symplectic form on $\mathcal{P}$ is denoted by Ω; the symplectic form on the reduced phase space $\widetilde{\mathcal{P}}$ is denoted by $\widetilde{\omega}$ (or simply ω).

$[a \sim b]$ *Classes of equivalence* The space of classes of equivalence of all objects a and b, belonging to some space S, identified under the relation "$\sim$" are typically denoted as $[a \sim b; a, b \in S]$. Sometimes the space S is clear from the context and omitted in the notation, $[a \sim b]$. Often the equivalence relation comes from a group G acting on S; then $[a \sim ga; a \in S, g \in G]$ is also called the space of G-orbits on S.

log *Logarithm* The natural logarithm (base e) is denoted by log.

l_g, r_g *Group translations* The left group translation $l_g : G \to G$, by a group element $g \in G$, is defined as $l_g(g') := gg'$. The right group translation $r_g : G \to G$, by g, is defined as $r_g(g') := g'g$.

Ad, ad *Adjoint maps* The adjoint action of a group element $g \in G$ on the group G is denoted by $\mathrm{Ad}_g : G \to G$. The adjoint action of a group element $g \in G$ on its Lie algebra $\mathfrak{g}$ is denoted by $\mathrm{ad}_g : \mathfrak{g} \to \mathfrak{g}$. The adjoint action of an algebra element $\xi \in \mathfrak{g}$ on the Lie algebra $\mathfrak{g}$ is denoted by $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$.

coad — Coadjoint maps — The coadjoint action of a group element $g \in G$ on its dual Lie algebra $\mathfrak{g}^*$ is denoted by $\mathrm{coad}_g : \mathfrak{g} \to \mathfrak{g}$. The coadjoint action of an algebra element $\xi \in \mathfrak{g}$ on the dual Lie algebra $\mathfrak{g}^*$ is denoted by $\mathrm{coad}_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$.

Ξ — Maurer-Cartan form — The Maurer-Cartan form is denoted by Ξ. It is a Lie algebra-valued 1-form on the Lie Group defined by $\Xi(X) = l_{g^{-1}*}X$, where $X \in T_gG$.

$\widetilde{\mathcal{P}}$, $\mathcal{Q}$ — Phase and configuration space — The (reduced) phase space is denoted by $\widetilde{\mathcal{P}}$. If it has a cotangent bundle structure, $\widetilde{\mathcal{P}} = T^*\mathcal{Q}$, the base space $\mathcal{Q}$ is referred to as the configuration space. For the causal diamond $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$.

ω, θ, H, X, $\{\,,\}$ — Symplectic structure and Poisson brackets — The symplectic 2-form on the phase space is denoted by ω. The symplectic potential, denoted by θ, is a 1-form satisfying $\omega = d\theta$. The phase space function H is associated with the Hamiltonian vector field X on phase space via $dH = -\imath_X\omega$. The Poisson brackets of two functions f and g is defined by $\{f, g\} := -\omega(X_f, X_g)$.

$\widetilde{G}$, G, $\widetilde{\mathfrak{g}}$, $\mathfrak{g}$, $\Gamma_{\widetilde{g}}$, δ_g, $\widetilde{\delta}_g$ — Groups and algebras of symmetry — The canonical group, acting transitively as symplectomorphisms of the phase space, will typically be denoted by $\widetilde{G}$. A group acting transitively on the configuration space, used as the basis for constructing the canonical group, will typically be called G. The Lie algebras of $\widetilde{G}$ and G will be denoted by $\widetilde{\mathfrak{g}}$ and $\mathfrak{g}$, respectively. The action of $\widetilde{G}$ on the phase space $\widetilde{\mathcal{P}}$ is denoted by $\Gamma_{\widetilde{g}}$; the action of G on the configuration space $\mathcal{Q}$ is typically denoted by δ_g, and its lift to the phase space (cotangent bundle) is $\widetilde{\delta}_g$. (Alternatively, sometimes we write simply gx for the action of g on x.)

$\rtimes$, $\oplus$ *Semidirect product and sum* The semidirect product of a group G and a group N, with respect to a homomorphism $\varphi : G \to \mathrm{Aut}(N)$, is denoted by $N \rtimes_\varphi G$ (or, omitting φ, simply $N \rtimes G$). It has the topology of $N \times G$, its element are denoted by $(n; g)$ and the product rule is $(n, g)(n', g') = (n\varphi_g(n); gg')$. The Lie algebra of this group corresponds to a semidirect sum of the respective Lie algebras, which is denoted by $\mathfrak{n} \oplus \mathfrak{g}$. (To remember the notation, note that the arrow in $\rtimes$ points from G to N, since G acts on N; for the algebra, note that $\oplus$ looks like a "rounded" arrow from $\mathfrak{g}$ to $\mathfrak{n}$.)

$\widehat{G}$, $\widehat{\mathfrak{g}}$ *Central extensions* The central extension of a group G by a real 2-cocycle is denoted by $\widehat{G}$. Its elements are $\widehat{g} = (g, r)$, where $g \in G$ and $r \in \mathbb{R}$. The product rule has the form $(g, r)(g', r') = (gg', r + r' + W(g, g'))$, where $W : G \times G \to \mathbb{R}$. (Note that (e, r) belongs to the center of $\widehat{G}$.) The central extension of a Lie algebra $\mathfrak{g}$ by a real 2-cocycle is denoted by $\widehat{\mathfrak{g}}$. Its elements are denoted by $\widehat{\xi} = \xi + r\widehat{c}$, where $\xi \in \mathfrak{g}$, $r \in \mathbb{R}$ and $\widehat{c}$ is the central element of $\widehat{\mathfrak{g}}$.

Diff$^+(S^1)$ *Circle diffeomorphisms* The group of (orientation-preserving) diffeomorphisms of

ψ, ϕ the circle is denoted by *Diff*$^+(S^1)$. Its elements are typically denoted by ψ :

$S^1 \to S^1$ or ϕ.

$\mathfrak{diff}(S^1)$, $\xi = \xi(\theta)\partial_\theta$, $\eta = \eta(\theta)\partial_\theta$ — *Algebra of circle diffeormorphisms* The Lie algebra of $\mathit{Diff}^+(S^1)$ is denoted by $\mathfrak{diff}(S^1)$. Its elements are identified with vector fields on S^1 and typically denoted by $\xi = \xi(\theta)\partial_\theta$ or $\eta = \eta(\theta)\partial_\theta$. The product rule is $[\xi, \eta] := [\xi, \eta]_{\mathfrak{diff}} := [\eta, \xi]_{S^1} := (\eta(\theta)\partial_\theta\xi(\theta) - \xi(\theta)\partial_\theta\eta(\theta))\,\partial_\theta$.

$\mathfrak{diff}^*(S^1)$, $\alpha = \alpha(\theta)d\theta^2$ — *Dual algebra of circle diffeormorphisms* The dual Lie algebra of $\mathit{Diff}^+(S^1)$ is denoted by $\mathfrak{diff}^*(S^1)$. Its elements can be identified with quadratic forms S^1 and typically denoted by $\alpha = \alpha(\theta)d\theta^2$ (or β). The pairing between $\mathfrak{diff}^*(S^1)$ and $\mathfrak{diff}(S^1)$ is $\alpha(\xi) := \int \alpha(\theta)d\theta^2(\xi(\theta)\partial_\theta) = \int d\theta\,\alpha(\theta)\xi(\theta)$.

$PSL(2,\mathbb{R})$, $\mathfrak{psl}(2,\mathbb{R})$, χ, υ — *Projective special linear group and algebra* The group of 2×2 real matrices S with unit determinant, where S is identified with $-S$, defines $PSL(2,\mathbb{R})$, whose elements are typically denoted by χ. Its Lie algebra is denoted by $\mathfrak{psl}(2,\mathbb{R})$, and here identified with the subalgebra of $\mathfrak{diff}(S^1)$ generated by the elements ∂_θ, $\sin\theta\partial_\theta$ and $\cos\theta\partial_\theta$. The elements of $\mathfrak{psl}(2,\mathbb{R})$ are sometimes denoted by υ. By exponentiating this subalgebra, $PSL(2,\mathbb{R})$ is realized as a subgroup of $\mathit{Diff}^+(S^1)$. (In Part I, it arises as the restriction to the boundary of the diffeomorphism component of the group of conformal isometries of the unit round disc.)

$\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, $[\psi] \in \mathcal{Q}$ — *Configuration space* The configuration space for the causal diamond is the quotient space $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$, and its points are denoted by $[\psi] = [\psi\chi]$.

Vira — $\underline{\text{Virasoro group}}$ The Virasoro group, $Vira = \widehat{Diff^+(S^1)}$, is a central extension of $Diff^+(S^1)$. Its elements are typically denoted by $\widehat{\psi} = (\psi, r)$, where $\psi \in Diff^+(S^1)$ and $r \in \mathbb{R}$. The identity element of *Vira* is denoted by $\widehat{I} = (I, 0)$, where I is the identity diffeomorphism of S^1.

$\widehat{\psi} = (\psi, r)$

$\widehat{I} = (I, 0)$

$\mathfrak{vira}$ — $\underline{\text{Virasoro algebra}}$ The Virasoro algebra, $\mathfrak{vira} = \widehat{\mathfrak{diff}(S^1)}$, is a central extension of $\mathfrak{diff}(S^1)$. Its elements are typically denoted by $\widehat{\xi} = \xi + x\widehat{c}$, where $\xi = \xi(\theta)\partial_\theta \in \mathfrak{diff}(S^1)$, $x \in \mathbb{R}$ and $\widehat{c}$ is the central element.

$\widehat{\xi} = \xi + x\widehat{c}$

$\widehat{\eta} = \eta + y\widehat{c}$

$\mathfrak{vira}^*$ — $\underline{\text{The dual of Virasoro algebra}}$ The dual of the Virasoro algebra is denoted by $\mathfrak{vira}^*$. Its elements are typically denoted by $\widetilde{\alpha} = \alpha + a\widetilde{c}$, where $\alpha = \alpha(\theta)d\theta^2 \in \mathfrak{diff}^*(S^1)$, $a \in \mathbb{R}$ and $\widetilde{c}$ is the dual of $\widehat{c}$ (i.e., $\widetilde{c}(\xi + x\widehat{c}) = x$).

$\widetilde{\alpha} = \alpha + a\widetilde{c}$

$\widetilde{\varepsilon} = d\theta^2 + \widetilde{c}$ — $\underline{\text{The representative of } \mathcal{Q}}$ In the realization of $\mathcal{Q} = Diff^+(S^1)/PSL(2, \mathbb{R})$ as a coadjoint orbit of *Vira*, $\widetilde{\varepsilon} := d\theta^2 + \widetilde{c} \in \mathfrak{vira}^*$ is point on the orbit (corresponding to $[I] \in \mathcal{Q}$).

$\underline{\psi} \in \underline{Diff^+(S^1)}$ — $\underline{\text{Universal covers}}$ The universal cover of a group G is denoted by an underline, $\underline{G}$. Accordingly, the universal cover of $Diff^+(S^1)$ is denoted by $\underline{Diff^+(S^1)}$, and its elements by $\underline{\psi}$. The universal cover of *Vira* is denoted by $\underline{Vira}$, and its elements by $\underline{\widehat{\psi}} = (\underline{\psi}, x)$.

$\underline{\widehat{\psi}} \in \underline{Vira}$

$S[\psi](\theta)$ — $\underline{\text{Schwarzian derivative}}$ The Schwarzian derivative maps circle diffeomorphisms ψ into real functions on the circle, $\psi \mapsto S[\psi](\theta)$.

exp *Lie Exponential* The Lie group exponential is typically denoted by $\exp : \mathfrak{g} \to G$. The group it refers to should be understood from the context, but typically will be the Virasoro group.

Ad_g
ad_g, ad_ξ

Adjoint maps The adjoint action of a group element $g \in G$ on the group G is denoted by $\mathrm{Ad}_g : G \to G$ and defined by $\mathrm{Ad}_g(g') := gg'g^{-1}$, where $g' \in G$. The adjoint action of a group element $g \in G$ on its Lie algebra $\mathfrak{g}$ is denoted by $\mathrm{ad}_g : \mathfrak{g} \to \mathfrak{g}$ and defined by $\mathrm{ad}_g := (\mathrm{Ad}_g)_*$, seen as map from T_eG to itself. The adjoint action of an algebra element $\xi \in \mathfrak{g}$ on the Lie algebra $\mathfrak{g}$ is denoted by $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ and defined by $\mathrm{ad}_\xi := \frac{d}{dt}\mathrm{ad}_{\exp(t\xi)}\big|_{t=0}$. (It is also true that $\mathrm{ad}_\xi \eta = [\xi, \eta]$.)

coad_g, coad_ξ

Coadjoint maps The coadjoint action of a group element $g \in G$ on its dual Lie algebra $\mathfrak{g}^*$ is denoted by $\mathrm{coad}_g : \mathfrak{g}^* \to \mathfrak{g}^*$ and defined by $\mathrm{coad}_g \alpha := \mathrm{Ad}^*_{g^{-1}}\alpha$, where $\alpha \in \mathfrak{g}^*$. The coadjoint action of an algebra element $\xi \in \mathfrak{g}$ on the dual Lie algebra $\mathfrak{g}^*$ is denoted by $\mathrm{coad}_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$ and defined by $\mathrm{coad}_\xi := \frac{d}{dt}\mathrm{coad}_{\exp(t\xi)}\big|_{t=0}$.

$\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$
$(\widehat{\eta}; \widehat{\phi})$
$(\eta + y\widehat{c}; (\phi, r))$

Canonical group The canonical group for the causal diamonds is taken to be $\widetilde{G} = \mathfrak{vira} \rtimes \mathit{Vira}$. The *Vira* factor act as "configuration translations" while the abelian factor, $\mathfrak{vira}$, act as "momentum translations". Elements of $\mathfrak{vira} \rtimes \mathit{Vira}$ are denoted as $(\widehat{\eta}; \widehat{\phi})$. The semi-colon separates the two components of $\widetilde{G}$, so the notation is more transparent when $\widehat{\phi}$ and $\widehat{\eta}$ are written explicitly, i.e., $(\widehat{\eta}; \widehat{\phi}) = (\eta + y\widehat{c}; (\phi, r))$.

$\widetilde{\mathfrak{g}} = \mathfrak{vira}^c \oplus \mathfrak{vira}$, $(\widehat{\eta};\widehat{\xi})$, $(\eta + y\widehat{c};\xi + x\widehat{c})$ — *Canonical algebra* The canonical algebra is the Lie algebra of $\widetilde{G}$, $\widetilde{\mathfrak{g}} = \mathfrak{vira}^c \oplus \mathfrak{vira}$, where $\mathfrak{vira}^c$ is the Lie algebra of abelian group $\mathfrak{vira}$ (i.e., $\mathfrak{vira}^c$ is isomorphic to $\mathfrak{vira}$ as vector space but has a commutative algebraic structure). Elements of $\mathfrak{vira}^c \oplus \mathfrak{vira}$ are denoted by $(\widehat{\eta};\widehat{\phi}) = (\eta + y\widehat{c};\xi + x\widehat{c})$.

L_n, R, K_n, T — *Fourier basis for* $\widetilde{\mathfrak{g}}$ A basis for $\widetilde{\mathfrak{g}}$ is defined as: $L_n := (0;e^{in\theta}\partial_\theta)$, $R := (0;\widehat{c})$, $K_n := (e^{in\theta}\partial_\theta;0)$ and $T := (\widehat{c};0)$. Note that L_n and R are generators of the *Vira* factor of $\widetilde{G}$ ("configuration translations"), and K_n and R are the generators of the $\mathfrak{vira}$ factor of $\widetilde{G}$ ("momentum translations"). Also, R and T belong to the center of $\widetilde{\mathfrak{g}}$.

$P_{\widehat{\xi}}$, $Q_{\widehat{\eta}}$, P_n, Q_n — *Canonical charges* The canonical charges associated with the *Vira* factor of $\widetilde{G}$ are called *momentum charges*, $P_{\widehat{\xi}} := H_{(0;\widehat{\xi})}$, and the canonical charges associated with the $\mathfrak{vira}$ factor of $\widetilde{G}$ are called *configuration charges* $Q_{\widehat{\eta}} := H_{(\widehat{\eta};0)}$. In the Fourier basis we have $P_n := P_{L_n}$ and $Q_n := Q_{K_n}$, while the central charges are realized as $P_R = 0$ and $Q_T = 1$.

BMS_3, $\mathfrak{bms}_3$ — *The Bondi-Metzner-Sachs group and algebra* The $\mathfrak{bms}_3$ algebra is a reduction of $\widetilde{\mathfrak{g}} = \mathfrak{vira}^c \oplus \mathfrak{vira}$ in which $R = (0;\widehat{c})$ is removed, and the BMS_3 group is its exponentiation. Note that $\mathfrak{vira} \rtimes \mathit{Vira}$ is a central extension of BMS_3.

π, q — *Projection maps* The projection map in the cotangent bundle is typically denoted by $\pi : T^*\mathcal{Q} \to \mathcal{Q}$. The quotient by $PSL(2,\mathbb{R})$ from $\mathit{Diff}^+(S^1)$ to $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ is denoted by q. By an abuse of notation, since $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ is identified with a coadjoint orbit of Virasoro, we also use q to denote the map from $\mathit{Diff}^{+}(S^1)$ (or Vira since the central element acts trivially) to $\mathcal{Q} \subset \mathfrak{vira}^*$ defined by $q(\psi) := \mathrm{coad}_{\psi}\widetilde{\varepsilon}$.

$\mathcal{O}$, $H_{\mathcal{O}}$ — Orbits and little groups: When a group G acts on a manifold $\mathcal{M}$, the orbit $\mathcal{O}$ of $x \in \mathcal{M}$ is the set of points gx for all $g \in G$. The subgroup H of G that fixes a point x (i.e., $gx = x$) is called the little group of x. The little group of points on the same orbit are the same (up to conjugation) so we denote it by $H_{\mathcal{O}}$ (and it is true that $\mathcal{O}$ is homeomorphic to $G/H_{\mathcal{O}}$). We are typically interested in coadjoint orbits of Virasoro, so $\mathcal{O} \subset \mathfrak{vira}^*$.

$F \hookrightarrow E \to B$ — Fiber bundles: A fiber bundle with total space E, base manifold M and fibers F is denoted by $F \hookrightarrow E \to B$. The second arrow corresponds to the bundle projection map from E to B, and the first (hooked) arrow indicates that F can be embedded into E as a fiber (although not uniquely).

$\mathcal{S} \hookrightarrow G \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$, $[\widehat{\psi}, \varsigma]$, $L_{\widehat{\psi}}$ — Associated vector bundle: The vector bundle associated with the principal bundle $H \hookrightarrow G \to \mathcal{O}$, where $\mathcal{O} = G/H$, with respect to the linear representation $\mathscr{U} : H \to \mathit{Aut}(\mathcal{S})$, is denoted by $\mathcal{S} \hookrightarrow G \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$. Its elements are classes of equivalence $[g, \varsigma] = [gh, \mathscr{U}(h^{-1})\varsigma]$, where $g \in G$, $h \in H$ and $\varsigma \in \mathcal{S}$. The action of G on $\mathcal{O}$ lifts to an action on the bundle defined by $L_{g'}[g, \varsigma] := [g'g; \varsigma]$. Typically we will consider the bundle $\mathcal{S} \hookrightarrow \mathit{Vira} \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$, whose elements will be denoted by $[\widehat{\psi}, \varsigma]$, and the lifted action is denoted by $L_{\widehat{\psi}}$ (or the one based on the universal cover of *Vira*, $\mathcal{S} \hookrightarrow \underline{Vira} \times_{\mathscr{U}} \mathcal{S} \rightarrow \mathcal{O}$).

$\mathcal{H}$ — $\Psi(\widetilde{\alpha})$, $\Psi([\psi])$ — *Hilbert space and wavefunctions* The Hilbert space $\mathcal{H}$ carries a (projective) irreducible unitary representation of the canonical group. In the wavefunction realization, $\mathcal{H}$ is identified with the space of sections on $\mathcal{S} \hookrightarrow G \times_{\mathscr{U}} \mathcal{S} \rightarrow \mathcal{O}$, whose elements are denoted by $\Psi(\widetilde{\alpha})$ where $\widetilde{\alpha} \in \mathcal{O}$. For the natural representation, based on $\mathcal{O} = \mathcal{Q}$, we can also use the notation $\Psi([\psi])$ where $[\psi] \in Diff^{+}(S^1)/PSL(2, \mathbb{R})$.

$\mathcal{S}$ — ς — *"Little" Hilbert space* In the wavefunction realization, the "little" Hilbert space $\mathcal{S}$ carries an irreducible unitary representation $\mathscr{U}$ of the little group H associated to a coadjoint orbit of Virasoro. Its elements will be denoted by ς, and they correspond to "internal states" of the wavefunction (i.e., like an intrinsic spin in quantum field theory). In the natural representation $\mathcal{S}$ carries a (projective) unitary irreducible representation of $PSL(2, \mathbb{R})$.

$\widehat{P}_n$, $\widehat{Q}_n$ — *Quantum operators* The quantum version of a classical observable, represented on the Hilbert space, is typically denoted by a $\widehat{\ }$ accent above the classical symbol. For example, the quantum operators associated with the canonical momentum and configuration charges, P_n and Q_n, are denoted by $\widehat{P}_n$ and $\widehat{Q}_n$, respectively.

$\mathcal{T}$ — *Twist* The twist of the corner of the diamond, as embedded in spacetime, is denoted by $\mathcal{T}$.

$[a \sim b]$ — *Classes of equivalence* The space of classes of equivalence of all objects a and

b, belonging to some space S, identified under the relation "$\sim$" are typically denoted as $[a \sim b; a, b \in S]$. Sometimes the space S is clear from the context and omitted in the notation, $[a \sim b]$. Often the equivalence relation comes from a group G acting on S; then $[a \sim ga; a \in S, g \in G]$ is also called the space of G-orbits on S.

l_g, r_g *Group translations* The left group translation $l_g : G \to G$, by a group element $g \in G$, is defined as $l_g(g') := gg'$. The right group translation $r_g : G \to G$, by g, is defined as $r_g(g') := g'g$.

Ξ *Maurer-Cartan form* The Maurer-Cartan form is denoted by Ξ. It is a Lie algebra-valued 1-form on the Lie Group defined by $\Xi(X) = l_{g^{-1}*}X$, where $X \in T_gG$. (In App. F we use an alternative definition based on the right group translation, $\Xi(X) = r_{g^{-1}*}X$.)

$\mathring{\sigma} \in \mathring{\mathfrak{diff}}^*(S^1)$ *"Sigma-circle"* The subspace of $\mathfrak{diff}^*(S^1)$ that annihilates $\mathfrak{psl}(2, \mathbb{R})$ is denoted by $\mathring{\mathfrak{diff}}^*(S^1)$ and its elements are typically denoted by $\mathring{\sigma}$.

$\widehat{\mathcal{S}}$, J *Partially-reduced phase space* In some sections we refer to the partially-reduced phase space $\widehat{\mathcal{S}} = \mathit{Diff}^+(S^1) \times \mathring{\mathfrak{diff}}^*(S^1)$. The projection map to the (fully) reduced phase space is denoted by $J : \widehat{\mathcal{S}} \to \widetilde{\mathcal{P}}$.

B The uniformization map

In this appendix we shall explain how to construct an explicit conformal map that transforms a generic Riemannian disc, (h_{ab}, D), into the reference Euclidean unit disc, $(\bar{h}_{ab}, D)$. In other words, we shall establish that the map considered in Sec. 6,

$$(\Psi, \Omega) \mapsto h_{ab} = \Psi_* \Omega \bar{h}_{ab} \tag{B.1}$$

is a surjection from $\mathrm{Diff}^+(D) \times C^\infty(D, \mathbb{R}^+)$ onto $\mathrm{Riem}(D)$. The construction also automatically implies that when the domain of the map is restricted to $\mathrm{Diff}^+(D) \times_\gamma C^\infty(D, \mathbb{R}^+)$ the surjection is onto $\mathrm{Riem}(D, \gamma)$. This is a particular form of the uniformization theorem, which says that every simply-connected Riemannian 2-manifold is conformally equivalent to the open unit disc, or the complex plane, or the Riemann sphere. In fact, the particular case we will consider is known as the *Riemann mapping theorem*. The explicit construction of such a map permits us to write down explicitly the projection map from the ADM phase space, described by metrics h_{ab} and (traceless) extrinsic curvatures σ^{ab}, into the reduced phase space $\widetilde{\mathcal{P}}$. More precisely, note that the first step of the reduction process discussed in Sec. 6, summarized in Fig. 4, consists of an *enlargement* of the phase space where (h_{ab}, σ^{ab}) is replaced by $(\Psi, \Omega, \bar{\sigma}^{ab})$, and the purpose of this appendix is to find a (Ψ, Ω) from a given h_{ab} (and after having done that, one would simply find $\bar{\sigma}^{ab}$ by computing $\Omega^2 \Psi^* \sigma^{ab}$). There are many (Ψ, Ω) that can uniformize any given h_{ab}, a consequence of certain ambiguities in the construction (in fact, as we know, these ambiguities should correspond to a $PSL(2,\mathbb{R})$ gauge).

Given a manifold (with boundary) D with the topology of a closed disc, let h_{ab} be a generic Riemannian metric on D and let $\bar{h}_{ab}$ be the metric that makes D a Euclidean unit disc, i.e., in polar coodinates $\{r,\theta\} \in [0,1] \times [0,2\pi)$ we have $\bar{h} = dr^2 + r^2 d\theta^2$. The construction of the map consist of two main steps:

(i) Use a Weyl transformation to "flatten" (h, D), so that it can be isometrically embedded as a region R of the complex plane (with its natural, Euclidean metric);

(ii) Construct an analytical map that deforms R into the unit complex disc $\mathbb{D} = \{z \in \mathbb{C}, |z| \leq 1\}$, which is isometric to $(\bar{h}, D)$.

B.1 Step (i): Flattening and embedding

For the first step we must find $\Gamma : D \to \mathbb{R}^+$ such that

$$\widehat{h}_{ab} = \Gamma h_{ab} \tag{B.2}$$

is flat. In two dimensions, flatness follows from requiring that the Ricci scalar vanishes,

$$0 = \widehat{R} = \frac{1}{\Gamma}\left(R - \nabla^2 \log \Gamma\right) \tag{B.3}$$

where R and $\nabla^2 = h^{ab}\nabla_a\nabla_b$ are respectively the Ricci scalar and Laplacian associated with h, and $\widehat{R}$ is associated with $\widehat{h}$. Therefore, Γ is solution of

$$\nabla^2 \log\Gamma = R \tag{B.4}$$

Note that the boundary conditions are not specified, so there are many solutions for this equations. In fact, note that this is the familiar Poisson equation for a "potential" $\log\Gamma$ and a "charge density" $-R$, so we know that for any choice of (Dirichlet) boundary values the equation has a unique (real) solution for $\log\Gamma$. For concreteness, we may (arbitrarily) choose $\Gamma|_{\partial D} = 1$. Now that we have a "flattened" disc, $(\widehat{h}, D)$, we construct an embedding isometry into the flat plane. To do so, choose an arbitrary point $p_0 \in D$ and an arbitrary orthonormal basis $\{e_1, e_2\}$, with respect to $\widehat{h}$, at p_0. Since $\widehat{h}$ is flat, this basis can be unambiguously extended to an orthonormal frame $\{e_1, e_2\}$ on the whole D by demanding that e_i is covariantly constant with respect to $\widehat{h}$,

$$\widehat{\nabla}_a (e_i)^b = 0 \tag{B.5}$$

and it matches with the basis chosen at p_0. Let $\{e^1, e^2\}$ denote the dual frame, satisfying $e^i(e_j) = \delta^i_j$ everywhere on D. We can define *Cartesian coordinates* on D by integrating these dual vectors along arbitrary curves starting from p_0, that is, we assign coordinates (x^1, x^2) to $p \in D$ via

$$x^i(p) := \int_{p_0}^{p} e^i \tag{B.6}$$

where the integral is along any curve in D joining p_0 and p. Note that since $\{e_i\}$ was defined though (B.5), the dual frame also consist of covariantly constant 1-forms, $\widehat{\nabla}_a(e^i)_b = 0$, which implies that $de^i = 0$ and therefore the choice of the curve joining p_0 to p does not affect the result of the integral. It follows that $e^i = dx^i$, and also that $\widehat{h} = \delta_{ij}e^i e^j = (dx^1)^2 + (dx^2)^2$, justifying the term "Cartesian" for these coordinates. In fact, this map $p \mapsto (x^1, x^2)$ is the isometric embedding into the Euclidean plane. Naturally, we can also express this map as an embedding in $\mathbb{C}$ by

$$\phi : D \to \mathbb{C}\,, \quad \phi(p) = x^1(p) + ix^2(p) \tag{B.7}$$

whose image shall be denoted by $R \subset \mathbb{C}$. This embedding is an isometry with respect to the Euclidean metric on $\mathbb{C}$,

$$w = \frac{dzd\bar{z} + d\bar{z}dz}{2} \tag{B.8}$$

that is, $w = \phi_*\widehat{h}$. Thus we have constructed a conformal tranformation from (h, D) to (w, R) where $w = \phi_*\Gamma h$.

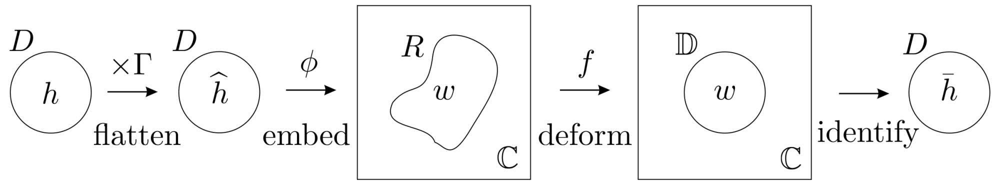


Figure 7: Illustration of the uniformization algorithm. The space (h, D) is flattened with a Weyl transformation, then isometrically embedded in the complex plane, then analytically deformed into the unit disc, and naturally identified with $(\bar{h}, D)$.

B.2 Step (ii): Deforming and identifying

Now we proceed to the second step, in which we analytically deform R into the unit complex disc $\mathbb{D} = \{z \in \mathbb{C}, |z| \leq 1\}$. That is, we want to construct an analytical map $f : R \to \mathbb{C}$ such that $\text{Im}(f) = \mathbb{D}$. The reason for requiring f to be analytic is so that it generates a conformal transformation on w, which can be seen as follows

$$f^* w = f^* \left(\frac{dz d\bar{z} + d\bar{z} dz}{2} \right) = \frac{d(z \circ f) d(\bar{z} \circ f) + d(\bar{z} \circ f) d(z \circ f)}{2} = |f'|^2 w \quad \text{(B.9)}$$

where in the last step we have used that $d(z \circ f) = (\partial f / \partial z) dz + (\partial f / \partial \bar{z}) d\bar{z} = f' dz$ and $d(\bar{z} \circ f) = \overline{d(z \circ f)} = \overline{f'} d\bar{z}$. Thus $w = f_* |f'|^2 w$, which corresponds to a conformal transformation from (w, R) to $(w, \mathbb{D})$. Since $(w, \mathbb{D})$ is naturally isometric to $(\bar{h}, D)$, we would have constructed a conformal transformation from (h, D) to $(\bar{h}, D)$ given by $\bar{h} = f_* |f'|^2 w = f_* |f'|^2 \phi_* \Gamma h = (f \circ \phi)_* (|f'|^2 \circ \phi) \Gamma h$. (See Fig. 7 for an illustration of the algorithm.) By inverting this transformation we would have the desired conformal map from $\bar{h}$ to h,

$$h = (f \circ \phi)_*^{-1} \left[\frac{1}{|f'|^2 (\Gamma \circ \phi^{-1})} \circ f^{-1} \right] \bar{h} \tag{B.10}$$

so, in the notation of (B.1), we identify $\Psi = (f \circ \phi)^{-1}$ and Ω as the function inside the square brackets. Thus it only remains to explain how to construct the map f. There is some freedom in the construction of this map, so for concreteness let us impose that $f(0) = 0$. With this choice, the point $p_0 \in D$ will end up at the origin of the unit disc. Moreover, we can try the following ansatz,

$$f(z) = z e^{g(z)} \tag{B.11}$$

where $g(z)$ is some analytic function. Both $u(z) := \mathrm{Re}[g(z)]$ and $v(z) := \mathrm{Im}[g(z)]$ are therefore harmonic (real) functions on $\mathbb{C}$. Now we impose that the boundary of R is mapped to the boundary of $\mathbb{D}$,

$$|f(z)| = 1 \text{ when } z \in \partial R \quad \Longrightarrow \quad u(z) = -\log|z| \text{ when } z \in \partial R \tag{B.12}$$

This fixes the value of u on ∂R, and since u is harmonic (with respect to w), there is a unique solution for u in the domain R. Also, $v(z)$ can be solved as

$$v(z) = \int_0^z dv = \int_0^z \left(\frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy \right) = \int_0^z \left(-\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy \right) \tag{B.13}$$

where $v(x,y) := v(x+iy)$ and we made the arbitrary choice $v(0) = 0$. With g fully determined in the domain R, f is also determined and the construction is therefore complete.

C Embedding the diamond in AdS_3

Since the diamond should be a region of AdS_3 it is interesting to have an explicit algorithm for, given any point in the phase space, constructing the corresponding diamond embedded in AdS_3. For this purpose, we make use of the fact that AdS_3 is a maximally-symmetric spacetime and, in particular, possesses a time-translation and a spatial-rotation symmetry.

C.1 General strategy

Consider the usual global coordinates on AdS_3, $\{t, r, \phi\}$, in which the metric takes the form

$$g = -\left(1 + \frac{r^2}{\ell_{AdS}^2}\right) dt^2 + \left(1 + \frac{r^2}{\ell_{AdS}^2}\right)^{-1} dr^2 + r^2 d\phi^2 \tag{C.1}$$

In this coordinate system, we can identify the time-translation symmetry as generated by $\xi^t = \partial/\partial t$ and the spatial-rotation as generated by $\xi^\phi = \partial/\partial\phi$. Observe that if we know these two Killing fields, in a generic connected patch of AdS_3, we can locally construct these coordinates as follows. Note that

$$dt_a = \frac{(\xi^t)_a}{(\xi^t)_b(\xi^t)^b} \tag{C.2}$$

so, up to an arbitrary assignment of time t_0 to a point x_0, the time coordinate t of any other point is given by

$$t(x) = t_0 + \int_{x_0}^{x} dt = t_0 + \int_0^1 ds \frac{(\xi^t)_a \dot{\gamma}^a}{(\xi^t)_b (\xi^t)^b} \tag{C.3}$$

where $\gamma(s)$ is any curve from x_0 at $s = 0$ to x at $s = 1$, and $\dot{\gamma}$ is the vector tangent to it. Similarly,

$$d\phi_a = \frac{(\xi^\phi)_a}{(\xi^\phi)_b (\xi^\phi)^b} \tag{C.4}$$

so we can assign angular coordinate ϕ to points as

$$\phi(x) = \phi_0 + \int_{x_0}^{x} d\phi = \phi_0 + \int_0^1 ds \frac{(\xi^\phi)_a \dot{\gamma}^a}{(\xi^\phi)_b (\xi^\phi)^b} \tag{C.5}$$

where ϕ_0 is an arbitrary angle coordinate for x_0, assumed not to be at the spatial origin (where the angle is undefined). The r coordinate of the point x can be extracted from the norm of either ξ^t or ξ^ϕ as

$$r(x) = \ell_{AdS} \sqrt{-(\xi^t)_a (\xi^t)^a - 1} = \sqrt{(\xi^\phi)_b (\xi^\phi)^b} \tag{C.6}$$

Of course, the condition that x_0 is not at the spatial origin is simply that $r(x_0) \neq 0$.

Another useful fact is that a Killing field is always uniquely determined from its value and first (anti-symmetric) derivative at any single point. This is so because a Killing field has no symmetric part for its first derivative, which implies that its second derivative is locally determined from its value as $\boldsymbol{\nabla}_a \boldsymbol{\nabla}_b \xi_c = \mathcal{R}_{cbad}\xi^d$,[87] and thus we can write the first-order system of equations

$$v^a \boldsymbol{\nabla}_a \xi_b = v^a \chi_{ab}$$
$$v^a \boldsymbol{\nabla}_a \chi_{bc} = v^a \mathcal{R}_{cba}{}^d \xi_d \tag{C.7}$$

where $\chi_{ab} := \nabla_a \xi_b = \nabla_{[a}\xi_{b]}$ and v^a is any vector field. If one takes $v = \dot{\gamma}$ as the tangent vector field along a curve γ, it is possible to integrate the equations from the initial values of ξ and χ. Using (3.39) we can rewrite the second equation as

$$v^a \boldsymbol{\nabla}_a \chi_{bc} = \frac{2}{\ell_{AdS}^2} v_{[b}\xi_{c]} \tag{C.8}$$

which is slightly simpler.

The final observation is that these equations can be restricted to any surface embedded in AdS_3, and can be solved as long as we know the induced metric h_{ab} and the extrinsic curvature K^{ab} on the surface. In particular, we can use the ADM data $(h_{ab}, \sigma^{ab}, \tau)$ on the CMCs to solve for the Killing fields and construct the desired coordinate system on it (which automatically provide the embedding of the CMCs into AdS_3). Consider a generic spacelike surface Σ embedded in AdS_3, with normal vector field n^a. The induced metric can be expressed as $h_{ab} = n_a n_b + g_{ab}$,[88] and the

[87] We use the boldface symbol $\boldsymbol{\nabla}$ to denote the three-dimensional covariant derivative associated with the AdS_3 metric g.

[88] More precisely, $h_a{}^b$ is the orthogonal projector to Σ, but h_{ab} coincides with the induced metric when rectricted to Σ.

extrinsic curvature is given by $K_{ab} = h_a{}^c \boldsymbol{\nabla}_c n_b$. Now we decompose the objects ξ and χ in their orthogonal ("hat") and tangent ("bar") components, as

$$\xi_a = g_{ab}\xi^b = (-n_a n_b + h_{ab})\xi^b = \widehat{\xi} n_a + \overline{\xi}_a \tag{C.9}$$

where

$$\begin{aligned} \widehat{\xi} &:= -n_a \xi^a \\ \overline{\xi}_a &:= h_{ab}\xi^b \end{aligned} \tag{C.10}$$

and

$$\chi_{ab} = g_a{}^{a'} g_a{}^{b'} \chi_{a'b'} = (-n_a n^{a'} + h_a{}^{a'})(-n_b n^{b'} + h_b{}^{b'})\chi_{a'b'} = 2n_{[a}\widehat{\chi}_{b]} + \overline{\chi}_{ab} \tag{C.11}$$

where

$$\begin{aligned} \widehat{\chi}_a &:= h_a{}^{a'} \chi_{a'b'} n^{b'} \\ \overline{\chi}_{ab} &:= h_a{}^{a'} h_b{}^{b'} \chi_{a'b'} \end{aligned} \tag{C.12}$$

If v is tangent to Σ, the variation of ξ becomes

$$\begin{aligned} v^a \boldsymbol{\nabla}_a \xi_b &= v^a \boldsymbol{\nabla}_a (\widehat{\xi} n_b + \overline{\xi}_b) \\ &= v^a \boldsymbol{\nabla}_a \widehat{\xi} n_b + \widehat{\xi} v^a (\boldsymbol{\nabla}_a n_b) + v^a \boldsymbol{\nabla}_a \overline{\xi}_b \\ &= v^a \nabla_a \widehat{\xi} n_b + \widehat{\xi} v^a K_{ab} + v^a \boldsymbol{\nabla}_a \overline{\xi}_b \end{aligned} \tag{C.13}$$

where in the last line it was used that $v^a \boldsymbol{\nabla}_a \widehat{\xi} = v^a \nabla_a \widehat{\xi}$, since the derivative of a scalar is independent of the connection, and also that $v^a K_{ab} = v^a h_a{}^c \boldsymbol{\nabla}_c n_b = v^a \boldsymbol{\nabla}_a n_b$, since v is tangent to Σ. For the last term we use the relation between the 3d derivative $\boldsymbol{\nabla}$ associated with g and the 2d derivative ∇ on Σ associated with h,

$$\nabla_a \overline{\xi}_b = h_a{}^{a'} h_b{}^{b'} \boldsymbol{\nabla}_{a'} \overline{\xi}_{b'} \tag{C.14}$$

which gives

$$v^a \nabla_a \overline{\xi}_b = v^a (n_b n^{b'} + g_b{}^{b'}) \boldsymbol{\nabla}_a \overline{\xi}_{b'} = v^a \boldsymbol{\nabla}_a \overline{\xi}_b - n_b v^a (\boldsymbol{\nabla}_a n^{b'}) \overline{\xi}_{b'} = v^a \boldsymbol{\nabla}_a \overline{\xi}_b - n_b v^a K_a{}^c \overline{\xi}_c \tag{C.15}$$

where we used that $v^a \boldsymbol{\nabla}_a (n^b \overline{\xi}_b) = 0$. Thus the first equation in (C.7) becomes

$$v^a \nabla_a \widehat{\xi} n_b + \widehat{\xi} v^a K_{ab} + v^a \nabla_a \overline{\xi}_b + n_b v^a K_a{}^c \overline{\xi}_c = v^a \chi_{ab} = -n_b v^a \widehat{\chi}_a + v^a \overline{\chi}_{ab} \tag{C.16}$$

which naturally decomposes into orthogonal and tangent parts,

$$v^a \nabla_a \widehat{\xi} = -v^a \widehat{\chi}_a - v^a K_a{}^c \overline{\xi}_c$$

$$v^a \nabla_a \overline{\xi}_b = -\widehat{\xi} v^a K_{ab} + v^a \overline{\chi}_{ab} \tag{C.17}$$

We now look at the variation of χ, which is given by

$$\begin{aligned} v^a \boldsymbol{\nabla}_a \chi_{bc} &= v^a \boldsymbol{\nabla}_a (2n_{[b} \widehat{\chi}_{c]} + \overline{\chi}_{bc}) \\ &= 2v^a \left(\boldsymbol{\nabla}_a n_{[b} \right) \widehat{\chi}_{c]} + 2n_{[b} v^a \boldsymbol{\nabla}_a \widehat{\chi}_{c]} + v^a \boldsymbol{\nabla}_a \overline{\chi}_{bc} \\ &= 2v^a K_{a[b} \widehat{\chi}_{c]} + 2n_{[b} v^a \left(\nabla_a \widehat{\chi}_{c]} + n_{c]} K_a{}^d \widehat{\chi}_d \right) + v^a \boldsymbol{\nabla}_a \overline{\chi}_{bc} \end{aligned} \tag{C.18}$$

where the middle term in the last line was manipulated analogously to (C.15). For the last term we have

$$\begin{aligned} v^a \nabla_a \overline{\chi}_{bc} &= h_b{}^{b'} h_c{}^{c'} v^a \boldsymbol{\nabla}_a \overline{\chi}_{b'c'} \\ &= (n_b n^{b'} + g_b{}^{b'})(n_c n^{c'} + g_c{}^{c'}) v^a \boldsymbol{\nabla}_a \overline{\chi}_{b'c'} \\ &= v^a \boldsymbol{\nabla}_a \overline{\chi}_{bc} - 2n_{[b} n^d v^a \boldsymbol{\nabla}_a \overline{\chi}_{c]d} \\ &= v^a \boldsymbol{\nabla}_a \overline{\chi}_{bc} + 2n_{[b} v^a (\boldsymbol{\nabla}_a n^d) \overline{\chi}_{c]d} \\ &= v^a \boldsymbol{\nabla}_a \overline{\chi}_{bc} + 2n_{[b} v^a K_a{}^d \overline{\chi}_{c]d} \end{aligned} \tag{C.19}$$

where in the forth line we used that $v^a \boldsymbol{\nabla}_a (n^d \overline{\chi}_{cd}) = 0$. So equation (C.8) becomes

$$2v^a K_{a[b}\widehat{\chi}_{c]} + 2n_{[b} v^a \nabla_a \widehat{\chi}_{c]} + v^a \nabla_a \overline{\chi}_{bc} - 2n_{[b} v^a K_a{}^d \overline{\chi}_{c]d} = -\frac{2}{\ell_{AdS}^2}\left(\widehat{\xi} n_{[b} + \overline{\xi}_{[b}\right) v_{c]} \tag{C.20}$$

which again can be decomposed into orthogonal and tangent parts,

$$\begin{aligned} v^a \nabla_a \widehat{\chi}_b &= v^a K_a{}^c \overline{\chi}_{bc} - \frac{1}{\ell_{AdS}^2}\widehat{\xi} v_b \\ v^a \nabla_a \overline{\chi}_{bc} &= -2v^a K_{a[b}\widehat{\chi}_{c]} - \frac{2}{\ell_{AdS}^2}\overline{\xi}_{[b} v_{c]} \end{aligned} \tag{C.21}$$

The system of equations (C.17) and (C.21) allows us to completely solve for ξ on the surface Σ given initial values for ξ and χ at a point of Σ.

This leads to the following algorithm for embedding the diamond in AdS_3,

1. Given a point $p \in \widetilde{\mathcal{P}}$ in the reduced phase space, consider any point $(\psi, \mathring{\sigma}) \in J^{-1}(p) \subset \widehat{\mathcal{S}}$ in the pre-image of p under J.

2. For a given time τ, solve the associated Lichnerowicz equation (6.19) for λ. Given any extension of ψ to a diffeomorphism Ψ of the disk, define $(h_{ab}, \sigma^{ab}) \in \mathcal{S}$ as in (6.16) and (6.17). On the CMC at time τ, with $K = -\tau$, we have $K^{ab} = \sigma^{ab} - \frac{1}{2}\tau h^{ab}$.

3. Choose an arbitrary point x_0 of the disk to be the origin of the AdS_3 coordinate system. We can always choose a boost such that $\partial/\partial t$ is orthogonal to the corresponding CMC at x_0. That is, $(\xi^t)^a = n^a$ and $(\chi^t)_{ab} = 0$ at x_0.

4. Solve the system of equations (C.17) and (C.21) on the disk along a curve from x_0 to a point $x_1 \in \partial D$ at the boundary using the initial data

$$\widehat{\xi}^t = 1\,, \quad (\overline{\xi}^t)_a = 0\,, \quad (\widehat{\chi}^t)_a = 0\,, \quad (\overline{\chi}^t)_{ab} = 0$$

5. Continue solving the equations along the whole boundary circumference, starting from x_1.

6. With $(\xi^t)^a$ determined along the boundary, compute the time coordinate of the points x of the boundary via

$$t(x) = t(x_1) - \int_{\theta(x_1)}^{\theta(x)} d\theta \frac{(\overline{\xi}^t)_a(\partial_\theta)^a}{(\widehat{\xi}^t)^2 - (\overline{\xi}^t)_b(\overline{\xi}^t)^b}$$

where θ is an arbitrary angular coordinate on the disk. Note that $t(x_1)$ can also be computed, given $t(x_0) = 0$, but it is unimportant.

7. At the coordinate origin, the angular Killing field vanishes and its first derivative is given by $(\chi^\phi)_{ab} = (e_1)_a(e_2)_b - (e_2)_a(e_1)_b$, where e_1 and e_2 are two orthogonal unit vectors orthogonal to $\partial/\partial t$. Note that, when restricted to the surface, it matches the volume element associated with the induced metric.

8. Solve the system of equations (C.17) and (C.21) on the disk along a curve

from x_0 to x_1 using the initial data

$$\widehat{\xi}^\phi = 0\,, \quad (\overline{\xi}^\phi)_a = 0\,, \quad (\widehat{\chi}^\phi)_a = 0\,, \quad (\overline{\chi}^\phi)_{ab} = \vartheta_{ab}$$

9. Continue solving the equations along the whole boundary circumference, starting from x_1.

10. With $(\xi^\phi)^a$ determined along the boundary, compute the angular coordinate of the points x of the boundary via

$$\phi(x) = \phi(x_1) + \int_{\theta(x_1)}^{\theta(x)} d\theta \frac{(\overline{\xi}^\phi)_a(\partial_\theta)^a}{(\overline{\xi}^\phi)_b(\overline{\xi}^\phi)^b - (\widehat{\xi}^\phi)^2}$$

Note that $\phi(x_1)$ can chosen arbitrarily, but it is unimportant.

11. Define the radial coordinate for points of the boundary as

$$r(x) = \ell_{AdS}\sqrt{(\widehat{\xi}^t)^2 - (\overline{\xi}^t)_a(\overline{\xi}^t)^a - 1} = \sqrt{(\overline{\xi}^\phi)_a(\overline{\xi}^\phi)^a - (\widehat{\xi}^\phi)^2}$$

12. With the boundary of one CMC successfully embedded in AdS_3, the diamond is determined as the domain of dependence of the interior of the boundary (where the "interior" is any spacelike disk attached to the boundary).

C.2 Numerical implementation and pictures

We have implemented the algorithm proposed above in *Mathematica* (Ver. 12). The document is named *Embedding_AdS3*. The program takes the input state $(\psi, \widehat{\sigma}) \in \widehat{\mathcal{S}}$ and plots the boundary loop (corner) of the diamond as embedded in AdS_3, in usual coordinates $\{t, x = r\cos\phi, y = r\sin\phi\}$.

We have plotted the corner of the diamond for some special states $(\psi, \widehat{\sigma})$. In Fig. 8 we have the case $\psi(\theta) = \theta$, $\widehat{\sigma}(\theta) = 0$, which corresponds to the "symmetric" diamond; note that the corner of the diamond is embedded as a planar circle. In Fig. 9 we have the case $\psi(\theta) = \theta + 0.15\sin(5\theta)$, $\widehat{\sigma}(\theta) = 0$, which corresponds to a "zero-momentum" diamond; note that the corner is still planar[89] but now oscillates spatially (5 times). In Fig. 10 we have the case $\psi(\theta) = \theta$, $\widehat{\sigma}(\theta) = 10\cos(5\theta)$, which corresponds to a "pure-momentum" diamond; note that the corner oscillates in a lightlike direction (5 times). Finally, in Fig. 11 we have the case $\psi(\theta) = \theta + 0.25(\cos(2\theta) - 1)$, $\widehat{\sigma}(\theta) = \sin(2\theta)$, which corresponds to a "spinning" diamond; although it may not be easy to see, the corner has a twist ($P_0 = 1.94$).

For artistic purposes, we have also developed a *Mathematica* program that produces mathematically accurate pictures of causal diamonds (in Minkowski space) given a (spatial, acausal) boundary loop. The future horizon is obtained by shooting geodesic light rays from the boundary loop, going inward and to the future; a light ray is terminated if it meets (or come sufficiently close, given a suitable thresh-

[89]The fact that "zero-momentum" diamonds are all planar can be seen from the time-reversal symmetry of the system.

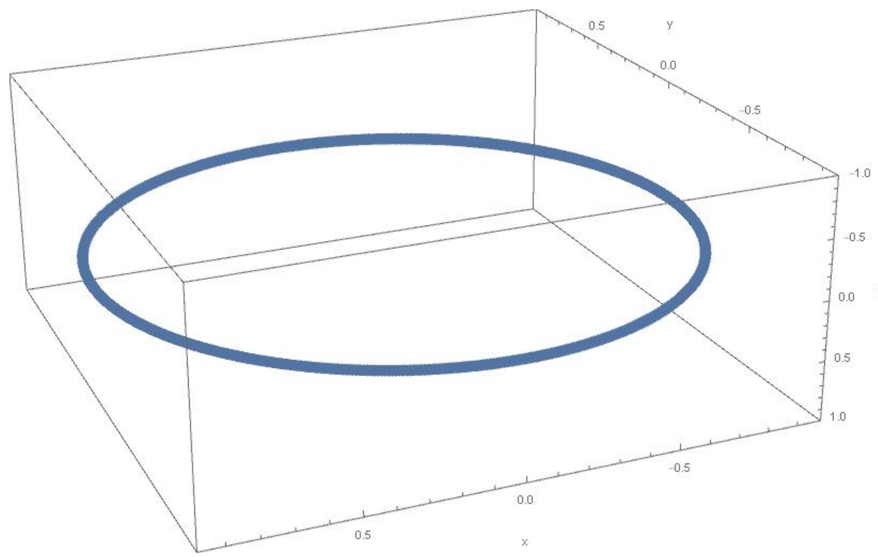


Figure 8: Corner of the "symmetric" diamond: $\psi(\theta) = \theta$, $\widehat{\sigma}(\theta) = 0$

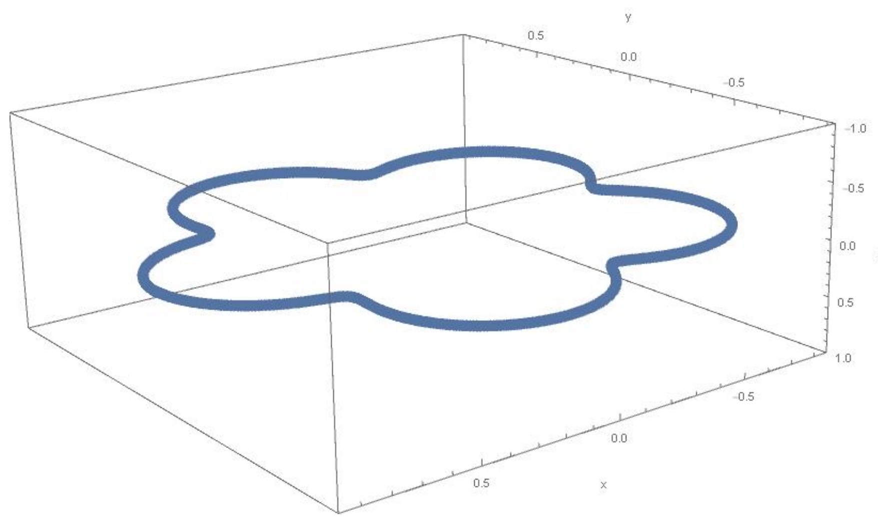


Figure 9: Corner of a "zero-momentum" diamond: $\psi(\theta) = \theta + 0.15\sin(5\theta)$, $\widehat{\sigma}(\theta) = 0$

old parameter) to the prolongation of any other light rays coming from the loop. The past horizon is produced in a similar fashion, using light rays going inwards and to the past

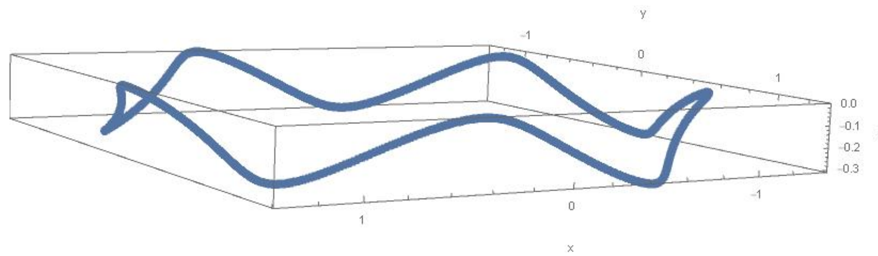


Figure 10: Corner of a "pure-momentum" diamond: $\psi(\theta) = \theta$, $\widehat{\sigma}(\theta) = 10\cos(5\theta)$

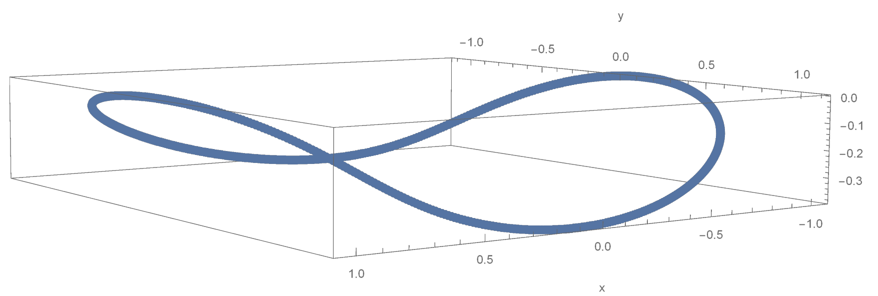


Figure 11: Corner of a "spinning" diamond ($P_0 = 1.94$): $\psi(\theta) = \theta + 0.25(\cos(2\theta) - 1)$, $\widehat{\sigma}(\theta) = \sin(2\theta)$

D Topology of $\mathcal{Q} = \mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$

We have argued that a more sophisticated method of quantization, such Isham's group-theoretic method, was necessary to handle the phase space $T^*(\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}))$ due to its non-linear structure. In this appendix we explain that this space does have a "trivial" topology (more precisely, contractible[90]), and it even admits a global chart of "Cartesian" coordinates (i.e., coordinates ranging from $-\infty$ to $+\infty$ that cover the entire space), but it appears that there is no preferred choice of coordinates due to the absence of an underlying linear structure.

Let us begin by quoting an important result in homotopy theory [172]. Given

[90] We thank E. Witten for bringing this to our knowledge.

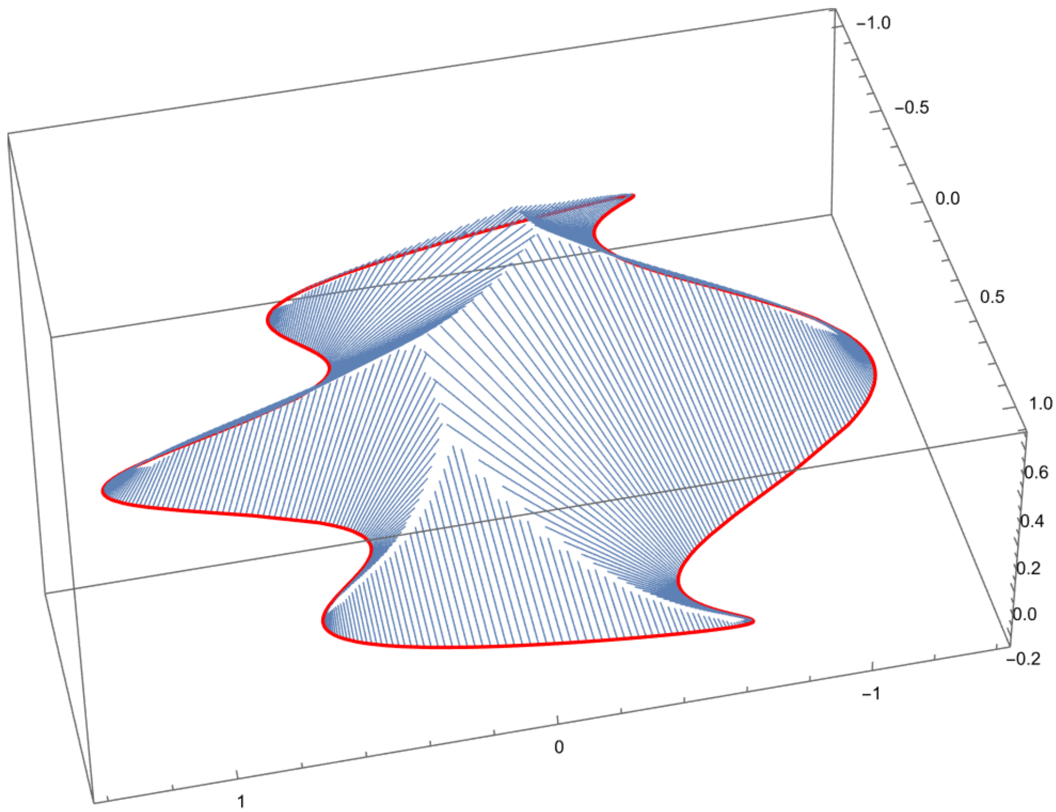


Figure 12: The future horizon of a causal diamond with a given boundary loop. The light rays are shown in blue and the loop in red.

a fiber bundle $F \hookrightarrow E \to M$, where the base space M is connected, there is an associated *exact* sequence of homotopy groups

$$\ldots \to \pi_n(F) \to \pi_n(E) \to \pi_n(M) \to \pi_{n-1}(F) \to \ldots \to \pi_0(F) \to 0 \tag{D.1}$$

where 0 denotes the trivial group. The maps featuring in this sequence are the following: $\pi_n(F) \to \pi_n(E)$ is the homomorphism induced from the fiber-into-bundle inclusion map $F \hookrightarrow E$ (while F is not canonically embedded as a fiber of E, any

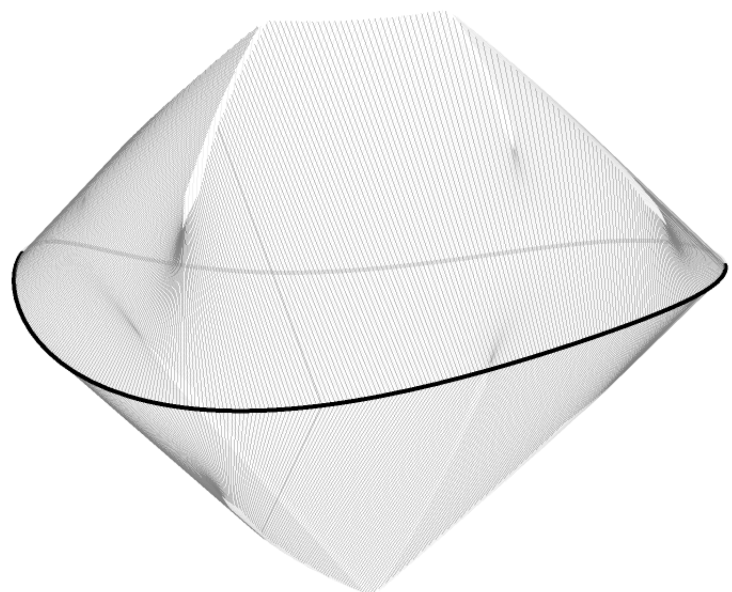


Figure 13: Another diamond with a finer spacing between the light rays.

such embedding defines the same homomorphism between the homotopy groups since the base space is connected); $\pi_n(E) \to \pi_n(M)$ is the homomorphism induced from the bundle projection map $E \to M$; and $\pi_n(M) \to \pi_{n-1}(F)$ is the boundary homomorphism (sometimes denoted by ∂) obtained by lifting (marked) n-spheres s on M to n-balls b on E, whose boundaries ∂b are $(n-1)$-spheres on the fiber F over the mark of s. This sequence is exact, meaning that the image of one map coincides with the kernel of the next, and all the maps are homomorphisms (with the possible exception of the last two since $\pi_0(F)$ may not be a group — however, as we will be interested in principal bundles, $\pi_0(F)$ has a group structure inherited from F).

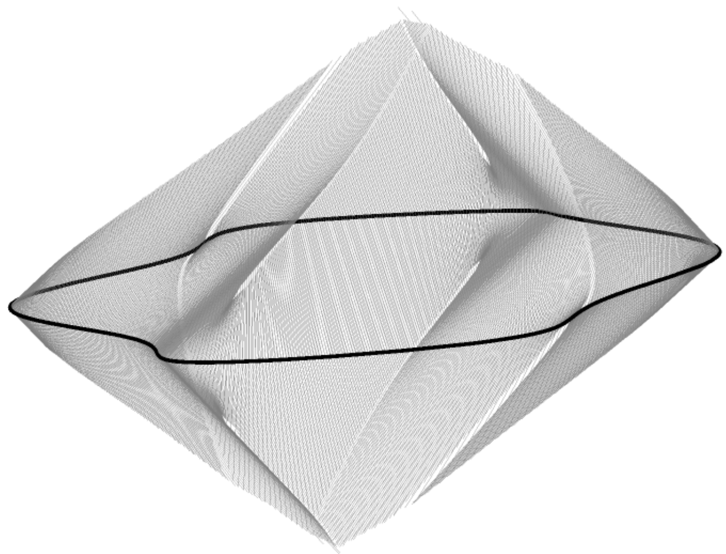


Figure 14: A diamond with a more complicated boundary loop.

Now we recall that $\mathit{Diff}^+(S^1)$ has fundamental group $\mathbb{Z}$. To see this, note that the group $\mathit{Diff}^+(\mathbb{R}; 2\pi)$ of (orientation-preserving) diffeomorphisms $f : \mathbb{R} \to \mathbb{R}$ of the real line ($f'(x) > 0$), satisfying the condition

$$f(x + 2\pi) = f(x) + 2\pi \tag{D.2}$$

is a (group) covering of $\mathit{Diff}^+(S^1)$. The projection map $\rho : \mathit{Diff}^+(\mathbb{R}; 2\pi) \to \mathit{Diff}^+(S^1)$ is defined from the identification $S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, where $\psi_f := \rho(f)$ acts on $\theta \in [0, 2\pi)$ as $\psi_f(\theta) = f(\theta) \bmod 2\pi$. The kernel of ρ, $\ker(\rho)$, which is a normal subgroup of $\mathit{Diff}^+(\mathbb{R}; 2\pi)$, consists of all translations by multiples of 2π, i.e., $f_n(x) = x + 2\pi n$ where $n \in \mathbb{Z}$, and thus it is isomorphic to the abelian group of integers $\mathbb{Z}$. Moreover, as to be expected, this (discrete) subgroup coincides with the center of $\mathit{Diff}^+(\mathbb{R};2\pi)$. Accordingly, we can think of $\mathit{Diff}^+(S^1)$ as the coset group

$$\mathit{Diff}^+(S^1) = \mathit{Diff}^+(\mathbb{R};2\pi)/\ker(\rho) = \mathit{Diff}^+(\mathbb{R};2\pi)/\mathbb{Z} \tag{D.3}$$

so that $\mathit{Diff}^+(\mathbb{R};2\pi)$ is a (principal) fiber bundle over the base $\mathit{Diff}^+(S^1)$ with fibers (structure group) $\mathbb{Z}$, which we write $\mathbb{Z} \hookrightarrow \mathit{Diff}^+(\mathbb{R};2\pi) \to \mathit{Diff}^+(S^1)$. Since $\mathit{Diff}^+(\mathbb{R};2\pi)$ is simply-connected, it is actually the universal cover of $\mathit{Diff}^+(S^1)$,

$$\underline{\mathit{Diff}^+(S^1)} = \mathit{Diff}^+(\mathbb{R};2\pi) \tag{D.4}$$

Inserting that $\pi_1(\mathit{Diff}^+(\mathbb{R};2\pi)) = 0$ and $\pi_0(\mathbb{Z}) = \mathbb{Z}$ into (D.1), we obtain the exact (sub)sequence

$$0 \to \pi_1(\mathit{Diff}^+(S^1)) \to \mathbb{Z} \to 0 \tag{D.5}$$

That $\mathbb{Z}$ maps to 0 implies that $\pi_1(\mathit{Diff}^+(S^1)) \to \mathbb{Z}$ is surjective, and that 0 maps to $\pi_1(\mathit{Diff}^+(S^1))$ implies that $\pi_1(\mathit{Diff}^+(S^1)) \to \mathbb{Z}$ is injective. Thus the fundamental group of $\mathit{Diff}^+(S^1)$ is

$$\pi_1(\mathit{Diff}^+(S^1)) = \mathbb{Z} \tag{D.6}$$

Evidently the non-contractible loops are associated with the $SO(2) \sim S^1$ subgroup of $\mathit{Diff}^+(S^1)$, and the homotopy classes are characterized by the winding number.

Finally, let us consider $\mathcal{Q} = \mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$.[91] One can anticipate that that $\mathcal{Q}$ is simply-connected as the $SO(2)$ inside $\mathit{Diff}^{+}(S^1)$ is "cancelled out" by the $SO(2)$ in $PSL(2,\mathbb{R}) \sim \mathbb{R}^2 \rtimes SO(2)$. In fact, a rotation by $2\pi n$ in $PSL(2,\mathbb{R})$ is mapped, is embedded as a subgroup of $\mathit{Diff}^{+}(S^1)$, to a rotation by $2\pi n$ in $\mathit{Diff}^{+}(S^1)$ when $PSL(2,\mathbb{R})$; therefore $\pi_1(PSL(2,\mathbb{R})) \to \pi_1(\mathit{Diff}^{+}(S^1))$ is an isomorphism. From (D.1), using that $\pi_0(PSL(2,\mathbb{R})) = 0$, we get

$$\mathbb{Z} \xrightarrow{\text{id}} \mathbb{Z} \to \pi_1\mathcal{Q} \to 0 \tag{D.7}$$

From the map $\pi_1\mathcal{Q} \to 0$ we conclude that $\mathbb{Z} \to \pi_1\mathcal{Q}$ must be a surjection, and from the fact that $\mathbb{Z} \to \mathbb{Z}$ is an identity (id) it follows that $\mathbb{Z} \to \pi_1\mathcal{Q}$ must be trivial. Therefore $\pi_1\mathcal{Q} = 0$. It is straightforward to show that all homotopy groups of $\mathcal{Q}$ are trivial, so $\mathcal{Q}$ is topologically contractible.

While not guaranteed that a contractible space is homeomorphic to a vector space, in this case one can construct global charts of "Cartesian coordinates" for $\mathcal{Q}$. A possible construction follows. First notice that given any two sets of three points on S^1, there is always a *psl* transformation that maps one set into the other. We can therefore use this property to "gauge-fix" the $PSL(2,\mathbb{R})$ action by restricting to diffeomorphisms that keep any three particular points fixed, e.g., $\theta = 0,\ 2\pi/3,\ 4\pi/3$. This is not a complete "gauge-fixing", since there is a $\mathbb{Z}_3$ subgroup of $SO(2) \subset PSL(2,\mathbb{R})$ that maps those three points to themselves. Therefore we can think of

[91] This space is a submanifold of the so-called *universal Teichmüller space*.

$\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$ as a trio of diffeomorphisms on a closed interval, up to cyclic permutation,

$$\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R}) = \mathit{Diff}^+([0,\pi])^3/\mathbb{Z}_3 \tag{D.8}$$

Also, $\mathit{Diff}^+([0,\pi])$ can be characterized by functions $f : [0,\pi] \to [0,\pi]$ satisfying $f(0) = 0$, $f(\pi) = \pi$ and $f'(x) > 0$. A manner to automatically enforce the last condition is to write $f'(x)$ as an exponential of a periodic function, which can be expanded in Fourier modes,

$$f'(x) = \kappa\, e^{\sum_n [a_n \sin(nx) + b_n \cos(nx)]} \tag{D.9}$$

where $n \geq 1$ and $\kappa := e^{b_0}$. Then $f(x) = \int_0^x dy f'(y)$ and the only constraint now is $f(\pi) = \int_0^\pi dy f'(y) = \pi$, which fixes κ in terms of a's and b's,

$$\kappa = \frac{\pi}{\int_0^\pi dx\, e^{\sum_n [a_n \sin(nx) + b_n \cos(nx)]}} \tag{D.10}$$

Thus, up to the discrete quotient by $\mathbb{Z}_3$, we can think of $\mathcal{Q}$ as being coordinatized by these six families of real parameters (i.e., for each of the three diffeomorphisms, families of a_n and b_n).

These coordinates can in principle be extended to the phase space, $\widetilde{\mathcal{P}} = T^*\mathcal{Q}$, by introducing the canonically conjugated pairs (i.e., the components of the 1-forms with respect to the coordinate basis, $p = p_i dq^i$). The symplectic form would locally take the Darboux format — the coordinates are not global because of the $\mathbb{Z}_3$

quotient, which would need to be addressed if one chooses to proceed with Dirac's quantization. It is important to stress, however, that these coordinates are arbitrary, and since it appears that there is no choice that is physically preferred, this approach cannot be used to justify a natural quantization.

Let us comment on a quasi-invariant measure on $\mathit{Diff}^{+}(S^1)/PSL(2,\mathbb{R})$.[92] First recall the classic *Wiener measure* on $C([0,1])$, the space of continuous real functions on the interval $[0,1]$, defined as follows [175, 176, 101]. Let $0 \leq x_1 < x_2 < \cdots < x_n \leq 1$ be any ordered set of n points in $[0,1]$, and $\{A_1, A_2, \ldots A_n\}$ be a collection of Borel sets on $\mathbb{R}$. A basis for the Borel σ-algebra on $C_0([0,1])$ consists of the sets $B(x_1, \ldots x_n; A_1, \ldots A_n)$ of functions $y \in C([0,1])$ satisfying $y(x_i) \in A_i$. The Wiener measure is defined on each basis element as

$$\mu_W\big[B(x_1,\ldots x_n; A_1,\ldots A_n)\big] := \int_{A_1} dy_1\, p(x_1;y_1) \int_{A_2} dy_2\, p(x_2-x_1;y_2-y_1)\cdots \int_{A_n} dy_n\, p(x_n-x_{n-1};y_n-y_{n-1}) \tag{D.11}$$

where

$$p(x;y) := \frac{e^{y^2/2x}}{\sqrt{2\pi x}} \tag{D.12}$$

As proposed by Shavgulidze [177], the Wiener measure can be pushed (for definition

[92] If G is a separable locally compact group and H is a closed subgroup, then the homogeneous space G/H admits a *unique* (up to equivalence) quasi-invariant measure with respect to G [173, 174]. (We are not aware if there are subtleties associated with infinite-dimensional groups.)

see footnote 93) to a measure μ_S on $\mathit{Diff}^+([0,1])$ via the correspondence:

$$
\begin{aligned}
C([0,1]) &\to \mathit{Diff}^+([0,1]) \\
y &\mapsto f(x) = \frac{\int_0^x dt\, e^{y(t)}}{\int_0^1 dt\, e^{y(t)}}
\end{aligned}
\tag{D.13}
$$

This measure is quasi-invariant with respect to the left-action by diffeomorphisms. We can define the product measure $\mu_S \otimes \mu_S \otimes \mu_S$ on $\mathit{Diff}^+([0,\pi])^3$ and push it under the Z_3 quotient to a measure μ on the configuration space, as given in (D.8), which is quasi-invariant with respect to the left-action by Virasoro (based on a similar result from [178, 101]).

Lastly, we remark that there exists a surprising, non-trivial symplectomorphism between $T^*\mathcal{Q}$, with the symplectic form associated with its cotangent bundle structure, and $\mathcal{Q} \times \mathcal{Q}$, with the symplectic structure that each factor $\mathcal{Q}$ inherits as a coadjoint orbit of Virasoro. This map is called the *Mess map*, first discovered in [179], and proven to be a symplectomorphism in [31]. (This map actually refers to the universal Teichmüller space, $\mathcal{T}(1)$, but modulo potential mathematical subtleties it descends to its submanifold $\mathit{Diff}^+(S^1)/PSL(2,\mathbb{R})$.) It is interesting that $\mathcal{Q} \times \mathcal{Q}$ is the most natural realization for the (reduced) phase space of vacuum asymptotically-AdS_3 gravity [29, 31, 180, 181]. However, the natural group to be quantized in that case is $\mathit{Vira} \times \mathit{Vira}$, which is of course the conformal group in two-dimensions (displaying a clear holographic aspect). Thus, while the underlying phase spaces may be the same, the structure of the symmetry groups $\mathit{Vira} \times \mathit{Vira}$

and $\mathfrak{vira} \rtimes Vira$ are distinct.

E Mackey's theory

In this appendix we outline the idea behind Mackey's theory of induced representations [182, 183, 173, 184]. In particular, we discuss how the concept of systems of imprimitivity can be used to prove irreducibility and exaustivity of the induced representations of semi-direct products of groups (at least in finite dimensions — but see [46]).

E.1 Induced representations

First let us recall how to construct a representation of a group G by inducing it from a representation of a subgroup H. Let $\mathscr{U} : H \to \mathrm{Aut}(\mathcal{S})$ denote a representation of H on a Hilbert space $\mathcal{S}$. From the bundle $H \to G \to G/H$, we can construct the associated bundle $\mathcal{S} \to G \times_{\mathscr{U}} \mathcal{S} \to G/H$ by "gluing" a copy of $\mathcal{S}$ at each point of G/H. More precisely, $G \times_{\mathscr{U}} \mathcal{S}$ is defined as the set of equivalence classes $[g, v] = [gh, \mathscr{U}_{h^{-1}} v]$, where $g \in G$, $h \in H$ and $v \in \mathcal{S}$, with projection map $\pi([g, v]) = gH \in G/H$. The induced representation from H to G, $U : G \to \mathrm{Aut}(\mathcal{H})$, acts on the space of cross sections, $\mathcal{H}$, of this associated bundle as

$$(U_g \Psi)(x) = \sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g(\Psi(g^{-1}x)) \tag{E.1}$$

where $x \in G/H$, $\Psi : G/H \to G \times_{\mathscr{U}} \mathcal{S}$ is a cross section ($\pi(\Psi(x)) = x$, for all x) and L_g is the G-action induced on the associated bundle defined by $L_g[g', v] := [gg', v]$. The Jacobian-like factor $\frac{d\mu_g}{d\mu}$, associated with a given quasi-invariant (Borel) measure μ on G/H, is the Radon-Nikodym derivative, with respect to μ, of the pushed measure μ_g through the group action, defined by $\mu_g[B] = \mu[g^{-1}B]$ for all Borel subsets B of G/H.[93] The inner product on $\mathcal{H}$ is defined by

$$\langle \Psi, \Psi' \rangle = \int_{G/H} d\mu(x) \left\langle\!\left\langle \Psi(x), \Psi'(x) \right\rangle\!\right\rangle \tag{E.2}$$

where the inner product inside the integral $\langle\!\langle\,,\rangle\!\rangle$ comes from the inner product on $\mathcal{S}$ $(\,,)$ and is defined by $\langle\!\langle \Psi(x), \Psi'(x) \rangle\!\rangle = \langle\!\langle [g, v(x)], [g, v'(x)] \rangle\!\rangle := (v(x), v'(x))$, where g is any element of G that projects to x under the quotient. Note that this is well-defined provided that $\mathscr{U}$ is unitary. Also note that U is unitary with respect to this

[93] More generally, given a measurable map $\rho : X \to Y$ and a measure μ on X, its push-forward of μ to Y is defined as $\rho_*\mu[B] = \mu[\rho^{-1}(B)]$, where B is any Borel subset of Y and ρ^{-1} denotes the pre-image under ρ.

inner product, as can be seen from

$$\begin{aligned}
\langle U_g\Psi, U_g\Psi'\rangle &= \int d\mu(x)\, \langle\!\langle U_g\Psi(x), U_g\Psi'(x)\rangle\!\rangle \\
&= \int d\mu(x) \left\langle\!\!\left\langle \sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g(\Psi(g^{-1}x)), \sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g(\Psi'(g^{-1}x)) \right\rangle\!\!\right\rangle \\
&= \int d\mu(x) \frac{d\mu_g}{d\mu}(x)\, \langle\!\langle L_g(\Psi(g^{-1}x)), L_g(\Psi'(g^{-1}x))\rangle\!\rangle \\
&= \int d\mu_g\, \langle\!\langle L_g(\Psi(g^{-1}x)), L_g(\Psi'(g^{-1}x))\rangle\!\rangle \\
&= \int d\mu_g\, \langle\!\langle \Psi(g^{-1}x), \Psi'(g^{-1}x)\rangle\!\rangle \\
&= \int d\mu(x)\, \langle\!\langle \Psi(x), \Psi'(x)\rangle\!\rangle \\
&= \langle \Psi, \Psi'\rangle
\end{aligned} \tag{E.3}$$

In the third line we just changed measures, $d\mu \frac{d\mu_g}{d\mu} = d\mu_g$, and in the sixth line we used the definition of the pushed measure, which implies $\int_B d\mu_g(x) f(x) = \int_{g^{-1}B} d\mu(x) f(gx)$ for any function f.

E.2 Systems of imprimitivity

Now we introduce the concept of systems of imprimitivity. But first let us define a *projection-valued measure* in analogy with the usual definition of a measure. Consider a manifold $\mathcal{M}$ and a Hilbert space $\mathcal{H}$. A projection-valued measure P is a map associating each Borel set B of $\mathcal{M}$ to an operator P_B on $\mathcal{H}$, satisfying:

(i) $P_{\mathcal{M}} = 1$ (i.e., identity operator on $\mathcal{H}$)

(ii) $P_{B\cap B'} = P_B P_{B'}$ for any Borel sets B and B'

(iii) $P_{B_1\cup B_2 \cup \cdots} = P_{B_1} + P_{B_2} + \cdots$ for any disjoint collection of Borel sets B_1, B_2, ...

Let G be a group that has an action on $\mathcal{M}$ and $U : G \to \mathrm{Aut}(\mathcal{H})$ be a representation of G on a Hilbert space $\mathcal{H}$. A *system of imprimitivity* for U based on $\mathcal{M}$ is a projection-valued measure P that transforms by conjugation under the group action, that is,

$$P_{gB} = U_g P_B U_{g^{-1}} \tag{E.4}$$

for all Borel sets B and $g \in G$. We shall refer to a system of system of imprimitivity by the pair (U, P), where the Hilbert space $\mathcal{H}$ (associated with U) and the manifold $\mathcal{M}$ (associated with P) are implicit.

Note that any representation of G induced from H is associated with a "canonical" system of imprimitivity based on G/H, with a natural projection-valued measure given by

$$P_B\Psi(x) := H_B(x)\Psi(x) \tag{E.5}$$

where $H_B : \mathcal{M} \to \mathbb{R}$ is the Heaviside function

$$H_B(x) = \begin{cases} 1\,, & x \in B \\ 0\,, & x \notin B \end{cases} \tag{E.6}$$

That this constitutes a system of imprimitivity can be shown as follows:

$$
\begin{aligned}
(U_g P_B U_{g^{-1}}\Psi)(x) &= \sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g\left[(P_B U_{g^{-1}}\Psi)(g^{-1}x)\right] \\
&= \sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g\left[H_B(g^{-1}x)(U_{g^{-1}}\Psi)(g^{-1}x)\right] \\
&= H_B(g^{-1}x)\sqrt{\frac{d\mu_g}{d\mu}(x)}\, L_g\left[\sqrt{\frac{d\mu_{g^{-1}}}{d\mu}(g^{-1}x)}\, L_{g^{-1}}(\Psi(x))\right] \\
&= H_B(g^{-1}x)L_g L_{g^{-1}}(\Psi(x)) \\
&= H_B(g^{-1}x)\Psi(x) = H_{gB}\Psi(x) \\
&= (P_{gB}\Psi)(x)
\end{aligned}
\tag{E.7}
$$

In the third line the Radon-Nikodym derivatives cancels because they combine into $\frac{d\mu_{gg^{-1}}}{d\mu} = \frac{d\mu}{d\mu} = 1$, following from the general formula $\frac{d\mu_{gg'}}{d\mu}(x) = \frac{d\mu_{g'}}{d\mu}(g^{-1}x)\frac{d\mu_g}{d\mu}(x)$. From the fourth to the fifth line we used $L_g L_{g'} = L_{gg'}$.

Given a group G, two systems of imprimitivity, (U, P) and (U', P'), based on the same manifold $\mathcal{M}$, are said to be equivalent if the representations, U and U', and the projection-valued measures, P and P', are related via conjugation by an isometry. More precisely, there exits an isometry $T : \mathcal{H} \to \mathcal{H}'$ such that

$$U' = TUT^{-1} \tag{E.8}$$

$$P' = TPT^{-1} \tag{E.9}$$

It is interesting to comment here on the measure introduced in the definition of the induced representation (E.1). We said that μ is a given quasi-invariant measure on G/H. In principle, each choice of μ could lead to a different representation. But this choice does not correspond to true arbitrariness because any two measures μ and μ' that are equivalent (i.e., have the same sets of measure zero) define equivalent representations. Moreover, they lead to equivalent systems of imprimitivity, as defined in (E.5). In fact, if G is locally compact, the choice is unique in the sense that all quasi-invariant measures on G/H are equivalent to each other [185]. This justifies the term "canonical" for the system of imprimitivity defined by (E.5). To see this equivalence, we note that any two equivalent measures are related by a Radon-Nikodym derivative, $d\mu(x) = \frac{d\mu}{d\mu'}(x)d\mu'(x)$, and it follows that

$$\frac{d\mu'_g}{d\mu'}(x) = \frac{\frac{d\mu}{d\mu'}(x)}{\frac{d\mu}{d\mu'}(g^{-1}x)}\frac{d\mu_g}{d\mu}(x) \tag{E.10}$$

The two representations U_g and U'_g defined as in (E.1), using μ and μ', act on the same Hilbert space $\mathcal{H}$. So the intertwiner T must be a unitary operator on $\mathcal{H}$. Let us try to define it as a multiplication by a positive function, $(T\Psi)(x) := \lambda(x)\Psi(x)$. We have,

$$\begin{aligned}(TU_gT^{-1}\Psi)(x) &= \lambda(x)(U_gT^{-1}\Psi)(x)\\ &= \lambda(x)\sqrt{\frac{d\mu_g}{d\mu}(x)}\,L_g(T^{-1}\Psi(g^{-1}x))\\ &= \frac{\lambda(x)}{\lambda(g^{-1}x)}\sqrt{\frac{d\mu_g}{d\mu}(x)}\,L_g(\Psi(g^{-1}x))\end{aligned} \tag{E.11}$$

so, with the choice $\lambda(x) = \sqrt{\frac{d\mu}{d\mu'}(x)}$, this gives $(U'_g\Psi)(x)$. It is also clear that $P'_B = TP_BT^{-1}$. Thus $(U, P) \sim (U', P')$.

Mackey's fundamental theorem is stated as follows:

The imprimitivity theorem Let G be a locally compact, separable group and H a closed subgroup of G. Let U be a unitary representation of G on a Hilbert space $\mathcal{H}$ and P be a system of imprimitivity for U on $\mathcal{M} = G/H$. Then there exists a unitary representation $\mathscr{U}$ of H on a Hilbert space $\mathcal{S}$ such that the canonical system of imprimitivity for the representation induced on G is equivalent to (U, P).

It follows that if every unitary representation of a certain group G can be associated with a *transitive* system of imprimitivity (i.e., based on a homogeneous space for G), then they all come from unitary representations of the corresponding "little groups" H by induction.

E.3 Semi-direct products

Here we consider groups of the form $A \rtimes G$, where A is a vector space[94] and G is a locally compact, separable group. The product rule is given by

$$(\alpha, g)(\alpha', g') = (\alpha + \delta_g\alpha', gg') \tag{E.12}$$

[94] Mackey's theory actually applies for abelian groups. When A is not a vector space, we just need to replace below the dual space A^* by the space of unitary characters $\mathrm{Char}(A)$.

where $\alpha \in A$, $g \in G$ and $\delta : G \to \mathrm{Aut}(A)$ is a left G-action on A (which need not be linear).

Note that a generic element of the group can be decomposed as

$$(\alpha, g) = (\alpha, e)(0, g) \tag{E.13}$$

where e is the identity element of G. Therefore, the operators representing $A \rtimes G$ on a Hilbert space $\mathcal{H}$ will factorize accordingly $U(\alpha, g) = U(\alpha, e)U(0, g)$. We can define $V(\alpha) := U(\alpha, e)$ and $D(g) := U(0, g)$, so that

$$U(\alpha, g) = V(\alpha)D(g) \tag{E.14}$$

Hence, in order to classify the representations of $A \rtimes G$, we can effectively study the representations of A and G separately, as we shall explain next.

Let us begin with A. We can define the self-adjoint generator $N(\alpha)$ by

$$V(\alpha) =: e^{-iN(\alpha)} \tag{E.15}$$

Since A is abelian, we have

$$V(\alpha)V(\alpha') = V(\alpha + \alpha') \quad \Rightarrow \quad e^{-iN(\alpha)}e^{-iN(\alpha')} = e^{-iN(\alpha+\alpha')} \tag{E.16}$$

and, since $N(\alpha)$ commutes with $N(\alpha')$, the product of exponentials is the exponential of the sum, so $N(\alpha) + N(\alpha') = N(\alpha + \alpha')$. Moreover, as $N((t + t')\alpha) = N(t\alpha) + N(t'\alpha)$ for any real numbers t and t', $N(t\alpha)$ must be linear in t, so $N(t\alpha) = tN(\alpha)$.[95] Thus,

$$N(t\alpha + t'\alpha') = tN(\alpha) + t'N(\alpha') \tag{E.17}$$

meaning that N is a linear map from A to a space of self-adjoint operators on $\mathcal{H}$. We can therefore think of N as an operator-valued element of the dual space A^*. Accordingly, a simultaneous basis of eigenvectors of $N(\alpha)$ can be labeled by elements $w \in A^*$ as

$$N(\alpha)|w\rangle = w(\alpha)|w\rangle \tag{E.18}$$

Note that the eigenvalues of $N(\alpha)$ need not be non-degenerate, so each $|w\rangle$ may be a vector in a Hilbert (sub)space $\mathcal{S}_w \subset \mathcal{H}$; and not all $w \in A^*$ need to be included in a given representation. What can we say about the relation between $\mathcal{S}_w$'s for different w? Note that, for every element (α', g') of the group,

$$(\alpha', g')(\alpha, e)(\alpha', g')^{-1} = (\delta_{g'}\alpha, e) \tag{E.19}$$

so A is a normal subgroup of $A \rtimes G$. At the level of the representation, we have

[95] More rigorously, define the operator-valued function of a real variable, $F(t) := N(t\alpha)$. As the representation of a Lie group is required to be smooth (i.e., a smooth homomorphism from the group into the space of linear operators), this function must be differentiable. Taking the s-derivative of $F(t+s) = F(t)+F(s)$, and evaluating at $s = 0$, gives $F'(t) = F'(0)$, where F' is the derivative of F with respect to its argument. Integrating from 0 to t gives $F(t) - F(0) = \int_0^t dt' F'(0) = tF'(0)$. But $F(0) = N(0) = 0$ since $U(0)$ is the identity, so $N(t\alpha) = tF'(0)$. Finally, evaluating this expression at $t = 1$ gives $F'(0) = N(\alpha)$, proving the result.

$U(\alpha', g')N(\alpha)U(\alpha', g')^{-1} = N(\delta_{g'}\alpha)$, so

$$N(\alpha)U(\alpha', g')|w\rangle = U(\alpha', g')N(\delta_{g'^{-1}}\alpha)|w\rangle = w(\delta_{g'^{-1}}\alpha)U(\alpha', g')|w\rangle = \widetilde{\delta}_{g'}w(\alpha)U(\alpha', g')|w\rangle \tag{E.20}$$

where $\widetilde{\delta}_{g'}w(\alpha) := \delta^*_{g'^{-1}}w(\alpha) = w(\delta_{g'^{-1}}\alpha)$ is the dual action of G in A^*. This means that $U(\alpha', g')$ maps $\mathcal{S}_w$ isometrically into $\mathcal{S}_{\widetilde{\delta}_{g'}w}$. In other words, $\mathcal{S}_w$ is isomorphic to $\mathcal{S}_{w'}$ as long as w and w' belong to the same G-orbit $\mathcal{O}$ in A^*. However, the isomorphism is not canonical for there is no unique group element that maps w into w'.

Considering the diagonalization above, the Hilbert space $\mathcal{H}$ is given by a direct "sum" of $\mathcal{S}_w$ over A^*,

$$\mathcal{H} = \int_{A^*} d\mu(w)\, \mathcal{S}_w \tag{E.21}$$

for some (Borel) measure μ on A^*. This is formalized by the Stone-Naimark-Ambrose-Godement theorem [174, 186]. A generic state $|\Psi\rangle \in \mathcal{H}$ can be expanded as

$$|\Psi\rangle = \int_{A^*} d\mu(w)|\Psi(w)\rangle \tag{E.22}$$

where $|\Psi(w)\rangle \in S_w \subset \mathcal{H}$. We can define a projector onto each $\mathcal{S}_w$ as

$$P_w|\Psi\rangle = |\Psi(w)\rangle \tag{E.23}$$

and, for each Borel set $B \subset A^*$, define

$$P_B = \int_B d\mu(w) P_w \tag{E.24}$$

The action on the each $\mathcal{S}_w$ can also be written, using the Heaviside function, as $P_B|\Psi(w)\rangle = H_B(w)|\Psi(w)\rangle$. Note that,

$$\begin{aligned} U(\alpha,g)P_B U(\alpha,g)^{-1}|\Psi\rangle &= \int_{A^*} d\mu(w)U(\alpha,g)P_B U(\alpha,g)^{-1}|\Psi(w)\rangle \\ &= \int_{A^*} d\mu(w)U(\alpha,g)H_B(\widetilde{\delta}_{g^{-1}}w)U(\alpha,g)^{-1}|\Psi(w)\rangle \\ &= \int_{A^*} d\mu(w)H_B(\widetilde{\delta}_{g^{-1}}w)|\Psi(w)\rangle \\ &= \int_{A^*} d\mu(w)H_{\widetilde{\delta}_g B}(w)|\Psi(w)\rangle = \int_{A^*} d\mu(w)P_{\widetilde{\delta}_g B}|\Psi(w)\rangle \\ &= P_{\widetilde{\delta}_g B}|\Psi\rangle \end{aligned} \tag{E.25}$$

In the second line we used that $U(\alpha,g)^{-1}|\Psi(w)\rangle \in \mathcal{S}_{\widetilde{\delta}_{g^{-1}}w}$. Hence, this is a projection-valued measure P on A^* that transforms by conjugation under G. This implies that (D, P) is a system of imprimitivity for G, based on A^*, where D is the restriction of U to G introduced in (E.14).

The system defined above is likely not transitive (unless G acts transitively on A^*), so we cannot apply the imprimitivity theorem yet. If the action is not transitive, this means that there must exist subsets Z of A^* which are invariant under the action of G, that is, $Z = \widetilde{\delta}_g Z$ for all $g \in G$. Assuming those are Borel sets, (E.25) implies that P_Z commutes with all operators $U(\alpha, g)$ of the representation. If we are interested in irreducible representations of $A \rtimes G$, then P_Z must be either 0 or 1 by Schur's lemma[96]. Of course, any two invariant sets Z and Z' must be disjoint and thus we must have $P_{Z \cup Z'} = P_Z + P_{Z'}$. Consequently, there must exist at most one invariant subset $Z \subset A^*$ that is equal to 1, all others must be 0. This means that the action of G on A^* is ergodic[97] with respect to the measure μ. Each such measure can lead to a different irreducible representation of $A \rtimes G$, so to exhaust all the possibilities we must classify all the inequivalent, quasi-invariant[98] measures on A^* with respect to which G acts ergodically. We shall consider this next.

Note that the orbits $\mathcal{O}$ in A^* are invariant subspaces. We have two possibilities for the measure:

(i) The measure is "concentrated" on a single orbit $\mathcal{O}$, i.e., $P_{\mathcal{O}} = 1$ for some orbit $\mathcal{O}$ (and 0 for all other orbits);

(ii) All orbits have measure zero. (This case is called *strictly ergodic*.)

The case (i) is the "simple" one. Since the measure on A^* is concentrated on a single orbit $\mathcal{O}$, it can be naturally restricted to $\mathcal{O}$, $\bar{\mu} := \mu|_{\mathcal{O}}$. Hence, the Hilbert space in (E.21) can be reduced to

$$\mathcal{H}_{\mathcal{O}} = \int_{\mathcal{O}} d\bar{\mu}(w)\, \mathcal{S}_w \tag{E.26}$$

[96]Schur's lemma implies that $P_Z = \lambda$, but $P_Z = P_{Z \cap Z} = P_Z P_Z$, so $\lambda(\lambda - 1) = 0$.

[97]A group action on a space is said to be *ergodic* if every invariant Borel subset is either of measure zero or the complement of a set of measure zero.

[98]Note that (E.25) implies that $P_B = 0 \Leftrightarrow P_{\widetilde{\delta}_g B} = 0$. From the definition of P_B, (E.24), this means that $\mu[B] = 0 \Leftrightarrow \mu[\widetilde{\delta}_g B] = 0$.

Note that this is the Hilbert space that we would have constructed by starting with a single vector $|\Psi(w)\rangle \in \mathcal{S}_w$ and acting with all operators $U(\alpha, g)$ on it. If we take P to be the projection-valued measure obtained by integrating over Borel sets of $\mathcal{O}$ with respect to $\bar{\mu}$, then (D, P) will be a transitive system of imprimitivity for G based on $\mathcal{O} \sim G/H$, where H is the "little group" associated with $\mathcal{O}$. Now we can apply the imprimitivity theorem to conclude that D must be induced from a unitary representation of H. This fully determines the irreducible representation $U(\alpha, g) = V(\alpha)D(g)$ on $\mathcal{H}_{\mathcal{O}}$. More concretely, realizing $\mathcal{H}_{\mathcal{O}}$ as sections of the bundle $\mathcal{S} \to G \times_{\mathscr{U}} \mathcal{S} \to \mathcal{O}$, where $\mathscr{U} : H \to \mathrm{Aut}(\mathcal{S})$ is the corresponding irreducible unitary representation of H on a Hilbert space $\mathcal{S}$ (isomorphic to any $\mathcal{S}_w$, $w \in \mathcal{O}$), then

$$\left(U(\alpha, g)\Psi\right)(w) = \left(V(\alpha)D(g)\Psi\right)(w) = e^{-iw(\alpha)}\sqrt{\frac{d\mu_g}{d\mu}(w)}\, L_g\Big(\Psi(\widetilde{\delta}_{g^{-1}}w)\Big) \tag{E.27}$$

where $w \in \mathcal{O} \subset A^*$ and μ is a quasi-invariant measure on $\mathcal{O}$.

To guarantee that there are no other representations, we need to ensure that case (ii) does not occur. One simple property which is sufficient is the following. A semi-direct product $A \rtimes G$ is said to be *regular* if the space of G-orbits in A^*, A^*/G, is measurable and there exists a measurable map $\zeta : A^*/G \to A^*$ associating to each orbit $\mathcal{O}$ a dual vector $w \in \mathcal{O}$. This map allows one to "pull-back" the measure μ in A^* to the orbits $\mathcal{O}$ in such a way that μ can be recovered by integration

[185, 182, 174]. More precisely, for any Borel set B in A^*,

$$\mu[B] = \int_{A^*/G} d\sigma(r)\, \bar{\mu}[B \cap \mathcal{O}_r] \tag{E.28}$$

where $r \in A^*/G$, σ is the measure on A^*/G, $\mathcal{O}_r \subset A^*$ is the orbit passing through $\zeta(r)$ and $\bar{\mu}$ is a measure on $\mathcal{O}_r$. In that case,

$$P_{A^*} = \int_{A^*} d\mu(w) P_w = \int_{A^*/G} d\sigma(r) \int_{\mathcal{O}_r} d\bar{\mu}(x) P_w = \int_{A^*/G} d\sigma(r) P_{\mathcal{O}_r} \tag{E.29}$$

where $x \in \mathcal{O}_r$ and $P_{\mathcal{O}_r}$ was defined by integrating P_w over $\mathcal{O}_r$ with $\bar{\mu}$. But if $P_{\mathcal{O}_r} = 0$ for all orbits (i.e., all r), then $P_{A^*} = 0$, which is a contradiction. Hence, the case (ii) does not occur in regular semi-direct products.

Another definition of regularity is the following. A semi-direct product $A \rtimes G$ is said to be *regular* if there is a countable family of Borel subsets Z_i of A^*, each a union of G-orbits, such that every orbit $\mathcal{O}$ is the intersection of a subfamily $Z_s^{\mathcal{O}}$ containing $\mathcal{O}$. The proof that this condition prevents case (ii) is similar to the above. If $P_{\mathcal{O}} = 0$ for every orbit, then

$$P[\cap_s Z_s^{\mathcal{O}}] = \prod_s P[Z_s^{\mathcal{O}}] = P[\mathcal{O}] = 0 \tag{E.30}$$

Since $P_{Z_i} = 0$ or 1, by Schur's lemma, then the formula above implies that at least one member of $\{Z_s^{\mathcal{O}}\}$, say $s = \bar{s}$, satisfies $P[Z_{\bar{s}}^{\mathcal{O}}] = 0$. Now consider the subfamily of Z_i composed of such members $Z_{\bar{s}}^{\mathcal{O}}$. Since they cover A^*, we would have $P_{A^*} = 0$, which is a contradiction.

Let us consider examples of a regular and a irregular semi-direct product. First, consider $\mathbb{R}^2 \rtimes SO(2)$ where $SO(2)$ acts as usual rotations. The dual action on $\mathbb{R}^{2*} \sim \mathbb{R}^2$ is also a rotation. The orbits decompose into circles (plus the origin), so $\mathbb{R}^2/SO(2) \sim \mathbb{R}^+ \cup \{0\}$. A possible choice of ζ is to associate the radius r of the orbit $C_r = \{x \in \mathbb{R}^2, |x| = r\}$ with the point it crosses the x-axis, i.e., $\zeta(r) = (r, 0)$. This is clearly a measurable map (given appropriate measures on $\mathbb{R}^+ \cup \{0\}$ and $\mathbb{R}^2$). For instance, if μ is the usual (Euclidean) measure on $\mathbb{R}^2$, the measure of a Borel set B can be decomposed as follows

$$\mu[B] = \int_0^\infty dr\, \bar{\mu}[B \cap C_r] \tag{E.31}$$

where $\bar{\mu}$ is the measure associated with $r d\theta$, with $0 \leq \theta < 2\pi$. Thus $\mathbb{R}^2 \rtimes SO(2)$ is regular. It is also regular with respect to the second definition because we can take Z_i to correspond to the family of all discs D_r, $r \in \mathbb{Q}$, together with their complements $\overline{D}_r = \mathbb{R}^2 - D_r$.

Second, let us consider $\mathbb{R}^2 \rtimes \mathbb{Z}$ where the cyclic group $\mathbb{Z}$ acts on $\mathbb{R}^2$ as rotations by irrational multiples of π, i.e., $(r, \theta) \mapsto (r, \theta + \pi\gamma)$ with γ irrational. The circles C_r are invariant under this action, but every orbit inside C_r is countable and thence have measure zero in the usual measure of the circle. The action is strictly ergodic for this measure. In particular, there exists irreducible unitary representations of

$\mathbb{R}^2 \rtimes \mathbb{Z}$ carried by wavefunctions on C_r, such as

$$U(\alpha, n)\Psi(x) = e^{-i\alpha\cdot x}\Psi(n^{-1}x) \tag{E.32}$$

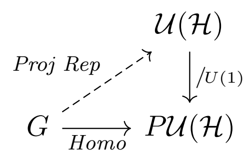
where $x \in C_r \subset \mathbb{R}^2$ and nx, with $n \in \mathbb{Z}$, denotes x rotated by $n\pi\gamma$. Since C_r is not an orbit of $\mathbb{Z}$, this is an "extra" representation of the group. This semi-direct product is not regular since no measurable map ζ exists, given that the space of orbits is $(S^1/\gamma\pi\mathbb{Z}) \times \mathbb{R}^+ \cup \{0\}$ is non-measurable.

F Projective representations

In this appendix we define projective representations of a Lie group, and explain their relationship with true representations of central extensions (by 2-cocycles) of the group. The intention is to offer a simple review of the subject, adapted to our notation. For details see [187] or, for a more informal discussion, [188]. Rigorously, the results here are only valid for finite-dimensional groups, even though we apply them to the quantization of causal diamonds, where the canonical group, $\mathfrak{vira} \rtimes \mathit{Vira}$, is infinite-dimensional. To simplify the language, we shall use "unirrep" as a short for "unitary irreducible representation".

F.1 Definition

If a physical symmetry acts on a Hilbert space $\mathcal{H}$ (i.e., so that it preserves all expectation values), then according to Wigner's theorem [189] the action must be unitary (linear) or anti-unitary (anti-linear). Let us consider the unitary case here, as the anti-unitary case is analogous. Since physical states are actually rays in $\mathcal{H}$, the unitary operators $U(g)$ and $e^{i\phi}U(g)$ implement the same physical transformation. Thus, the physically relevant space of transformations is the quotient $P\mathcal{U}(\mathcal{H}) := \mathcal{U}(\mathcal{H})/U(1)$ of unitary operators on $\mathcal{H}$ modulo a phase. In this manner, a group G of physical symmetries is realized in quantum mechanics as a homomorphism from G into $P\mathcal{U}(\mathcal{H})$, that is, $[U(g)][U(g')] = [U(gg')]$. Such a homomorphism is called a *projective (unitary) representation of G on $\mathcal{H}$.*



Given some (arbitrary, local) association $g \mapsto U(g)$, we have

$$U(g)U(g') = e^{i\phi(g,g')}U(gg') \tag{F.1}$$

for some real function $\phi : G \times G \to \mathbb{R}$. From associativity, ϕ must satisfy

$$\phi(g, g') + \phi(gg', g'') = \phi(g', g'') + \phi(g, g'g'') \tag{F.2}$$

which is called the *cocyle condition.* Applying this condition for $g' = e$ we see that $\phi(g, e) = \phi(e, g'')$, implying that $\phi(g, e) = \phi(e, g) = \phi(e, e)$. Without loss of generality, let us take $\phi(g, e) = \phi(e, g) = 0$ so that $U(e) = 1$. Note that if we choose another association, $g \mapsto U'_g = U_g e^{i\alpha(g)}$, where $\alpha : G \to \mathbb{R}$ is some function, we get a redefinition of ϕ given by $\phi'(g, g') = \phi(g, g') + \alpha(g) + \alpha(g') - \alpha(gg')$. This new ϕ automatically satisfies the cocycle condition, provided that the old ϕ does, and clearly these ϕ's define equivalent projective representations. So we have a *cohomology*, where a generic ϕ satisfying the cocycle condition is called a 2-cocycle, a 2-cocycle with the particular form $\phi(g, g') = \alpha(gg') - \alpha(g) - \alpha(g')$ is called a 2-coboundary, and the set of equivalence classes in which two 2-cocycles that differ by a 2-coboundary are identified is called a 2-cohomology.

F.2 Central extension by a 2-cocycle

Given a function $\phi : G \times G \to \mathbb{R}$ satisfying the cocycle condition (F.2), and for convenience $\phi(e, e) = 1$, we define the group extension

$$G_\phi := G \times_\phi \mathbb{R} \tag{F.3}$$

by the product rule

$$(g, r)(g', r') = (gg', r + r' + \phi(g, g')) \tag{F.4}$$

The cocycle condition is necessary to ensure associativity of G_ϕ. Note that (e, r) is in the center. A generic element of the extended group can be factorized as $(g, r) = (e, r)(g, 0)$, so a (true) representation of G_ϕ will satisfy $U(g, r) = U(e, r)U(g, 0)$, and we can define $V(r) := U(e, r)$ and $D(g) := U(g, 0)$. Due to Schur's lemma, in a (complex) irreducible representation the central elements are represented as multiples of the identity. Therefore, $V(r)$ forms a unitary irreducible representation of $\mathbb{R}$, and consequently $V(r) = e^{i\alpha r}$ for some $\alpha \in \mathbb{R}$. Thus,

$$U(g, r) = e^{i\alpha r} D(g) \tag{F.5}$$

Clearly, $D(g)D(g') = e^{i\alpha\phi(g,g')} D_(gg')$, so D is a projective representation of G associated with the phase $\alpha\phi$. But if ϕ satisfies the cocycle condition, then so does $\alpha\phi$. We thus conclude that every (true) unitary irreducible representation of G_ϕ corresponds to a projective unitary irreducible representation of G. Also, notice that if ϕ' and ϕ are cohomologous (i.e., differ by a coboundary), then $G_{\phi'}$ and G_ϕ are homomorphic, so it follows that equivalent (true) unirreps of G_ϕ define equivalent projective unirreps of G.

The Lie algebra of G_ϕ, denoted by

$$\mathfrak{g}_\varphi := \mathfrak{g} \oplus_\varphi \mathbb{R} \tag{F.6}$$

has the product structure

$$[(\xi, a), (\xi', a')] = ([\xi, \xi], \varphi(\xi, \xi')) \tag{F.7}$$

where $\varphi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is derivative of ϕ with respect to both of its arguments (i.e., the push-forward on both the first and second entries of ϕ). The function φ is anti-symmetric, bilinear and satisfies the cocycle condition

$$\varphi(\xi,[\xi',\xi'']) + \varphi(\xi',[\xi'',\xi]) + \varphi(\xi'',[\xi,\xi']) = 0 \tag{F.8}$$

Note that this cocycle condition can either be seen as following from the cocycle condition (F.2) for ϕ, or as following directly from the Jacobi identity. There is also an analogous cohomology for the algebra, where the coboundaries (i.e., the trivial elements) are those φ's with form $\varphi(\xi,\xi') = f([\xi,\xi'])$, where f is any real linear function on $\mathfrak{g}$. Naturally, a central extension of G is only possible if its algebra $\mathfrak{g}$ admits a central extension by 2-cocycles. If the algebra provides an obstruction to this construction, by not admitting (non-trivial) central extensions, then the group cannot be extended in this manner.

F.3 Central extension by a discrete group

It may still be possible to further centrally extend G_ϕ by a *discrete* abelian group, that is, to "unwrap" it to some covering groups. True representations of the covering groups also define projective representations of the group. The point is that the exponentiation of any unirrep of $\mathfrak{g}_\varphi$ defines a unirrep of the universal cover of G_ϕ and, at the same time, a projective unirrep of G, as we will see in this section. In this subsection we shall omit the labels ϕ and φ, since we are concerned only with discrete extension (that is, assume that G has already been extended by a 2-cocycle, if possible). We remark that the approach here is inspired by [190], particularly App. 2.B, but rephrasing it in a more geometrical language.

Let us recall how "exponentiate" a representation of the algebra $\mathfrak{g}$, $A : \mathfrak{g} \to \mathrm{SelfAdj}(\mathcal{H})$ into a representation of the (universal cover) of the group. Consider a curve $\gamma : [0,1] \to G$ starting at e and ending at g, and let $\dot{\gamma}(t)$ be the vector tangent to it at the parameter t. Since A acts on the algebra, we introduce the (right-invariant) Cartan-Maurer form, Ξ, to map $\dot{\gamma}(t)$ to $\mathfrak{g}$.[99] Consider the following equation

$$\frac{d}{dt}U(\gamma(t)) = -iA(\Xi(\dot{\gamma}(t)))U(\gamma(t)) \tag{F.9}$$

with initial condition $U(\gamma(0)) = U(e) = 1$, and define $U(\gamma) := U(\gamma(1))$. The solution is the path-ordered exponential

$$U(\gamma) = \mathcal{P}\mathrm{exp}\left(-i\int_0^1 dtA(\Xi(\dot{\gamma}(t)))\right) \tag{F.10}$$

This can be interpreted as a parallel transportation in the (trivial) principal bundle

$$\mathcal{U}(\mathcal{H}) \hookrightarrow \mathcal{U}(\mathcal{H}) \times G \to G \tag{F.11}$$

with fibers $\mathcal{U}(\mathcal{H})$, base manifold G, projection map $(U,g) \mapsto g$, and the group

[99] The *right-invariant* Cartan-Maurer form uses the right-translation, $r_g(g') := g'g$, to map vectors to the identity, i.e., if $v \in T_gG$ then $\Xi(v) := r_{g^{-1}*}v$. Note that this is not the standard definition, which uses left-translations instead. Our choice is better adapted to the conventions of this section, in particular because of the introduction of the complex unit in the exponential, so that A is self-adjoint instead of anti-self-adjoint.

structure is $\mathcal{U}(\mathcal{H})$ acting from the right as $(U,g)U' = (UU',g)$. Accordingly, $-A\Xi$ is interpreted as a (representative of a) connection on this bundle, since it takes a tangent vector at each point on the base and returns a value in the algebra of the structure group. In this way, the formula above is expressing that $U(\gamma)$ is the parallel transport of 1 along γ (i.e., $\gamma(t) \mapsto (U(\gamma(t)), \gamma(t))$ is a horizontal lift).

Now we prove that this connection is flat. Denote the connection representative by $\omega = -A\Xi$. (This representative is associated with the trivial section, $g \mapsto (1,g)$.) The curvature is given by

$$F(X,Y) = d\omega(X,Y) + \omega(X) \star \omega(Y) \tag{F.12}$$

where $\star$ denotes the product in the underlying algebra of the bundle, SelfAdj$(\mathcal{H})$, defined by $\star = -i[\,,]$. Since this is a tensor, its value at some point g depends only on the local values of the fields X and Y, so we might choose them to be right-invariant, i.e., let $X = r_{g*}\xi$ and $Y = r_{g*}\eta$. Thus, $\omega(X) = -A(\xi)$ and $\omega(Y) = -A(\eta)$ are constant functions on G. The first term gives $d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) = A\Xi([X,Y]) = -A([\xi,\eta])$, where we used $[r_{g*}\xi, r_{g*}\eta] = -r_{g*}[\xi,\eta]$. The second term gives $\omega(X) \star \omega(Y) = A(\xi) \star A(\eta) = A([\xi,\eta])$, where we used that A is a homomorphism from $\mathfrak{g}$ to SelfAdj$(\mathcal{H})$. Thus, the two terms cancel and we get $F = 0$. This implies that $U(\gamma)$ depends only on the homotopy class of the curve γ joining e and $\gamma(1)$.

Here we prove that, given two curves γ and γ' starting from e, we have $U(\gamma)U(\gamma') =$

$U(\gamma\gamma')$, where the product of curves is defined as: go along $\gamma'(t)$ with double speed for half the time, and then continue along $\gamma(t)\gamma'(1)$ with double speed until $\gamma(1)\gamma'(1)$. More precisely,

$$\gamma\gamma'(t) := \begin{cases} \gamma'(2t) & \text{for } 0 \leq t < \frac{1}{2} \\ \gamma(2t-1)\gamma'(1) & \text{for } \frac{1}{2} < t \leq 1 \end{cases} \tag{F.13}$$

Denote the vector tangent to $\gamma\gamma'(t)$ by $v(t)$. For the first half of the integration, we have $v(t) = 2\dot{\gamma}'(2t)$. By changing variables, $s = 2t$, that we get $U(t = \frac{1}{2}) = U(\gamma'(1)) = U(\gamma')$. For the second half, note that $v(t) = r_{g'*}(2\dot{\gamma}(2t-1))$, so that $\Xi(v(t)) = 2\Xi(\dot{\gamma}(2t-1))$. Multiplying the equation on both sides (from the right) by $U(\gamma')$, we get

$$\frac{d}{dt}\left(U(\gamma\gamma'(t))U(\gamma')^{-1}\right) = -iA\left(2\Xi(\dot{\gamma}(2t-1))\right)\left(U(\gamma\gamma'(t))U(\gamma')^{-1}\right) \tag{F.14}$$

so changing the variables to $s = 2t-1$, the solution (at $t = 1$) is just $U(\gamma\gamma')U(\gamma')^{-1} = U(\gamma)$. That is,

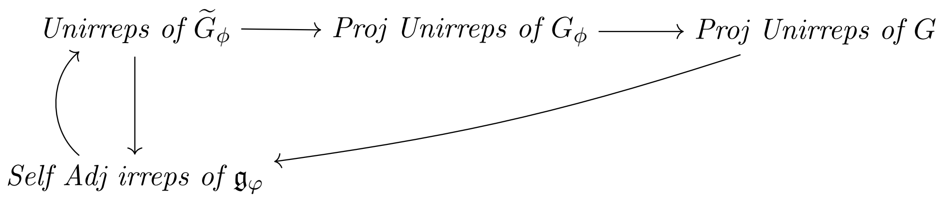
$$U(\gamma)U(\gamma') = U(\gamma\gamma') \tag{F.15}$$

revealing that U is a homomorphism from the space of homotopy classes of curves on G (starting at e) to unitary operators on $\mathcal{H}$. But, if G is connected, the space of homotopy classes of curves on G (starting at e) is precisely the universal cover, $\widetilde{G}$, of G. Therefore U defines a unitary representation of $\widetilde{G}$ on $\mathcal{H}$. In fact, as any unitary representation of $\widetilde{G}$ defines a self-adjoint representation of $\mathfrak{g}$, we conclude that there is a one-to-one correspondence between unitary representations of $\widetilde{G}$ and self-adjoint representations of $\mathfrak{g}$.

Finally, let us see how irreps of $\widetilde{G}$ descend to projective irreps of G. Denote a curve on G going from e to g by γ_g. We can always write $\gamma_g\gamma_{g'} := \eta(g,g')\gamma_{gg'}$, where η is a curve going from e to e (i.e., an element of the fundamental group $\pi_e(G)$). Note that η is an element on the fiber over e in $\widetilde{G}$, which forms a normal subgroup of $\widetilde{G}$ and, being discrete, must be in the center of $\widetilde{G}$. It follows, by considering the triple product $\gamma_g\gamma_{g'}\gamma_{g''}$, that η also satisfies a cocycle condition, $\eta(g,g')\eta(gg',g'') = \eta(g',g'')\eta(g,g'g'')$. Now consider $U(\gamma_g)U(\gamma_g') = U(\gamma_g\gamma_{g'}) = U(\eta(g,g'))U(\gamma_{gg'})$. For unirreps, $U(\eta(g,g')) = e^{i\phi(g,g')}$. The cocycle condition for η implies that ϕ satisfies the cocycle condition (F.2). Thus, defining $D(g) := U(\gamma_g)$, we have $D(g)D(g') = e^{i\phi(g,g')}D(gg')$.

F.4 A diagram and an example

We have seen in Sec. F.3 that self-adjoint irreps of $\mathfrak{g}_\varphi$ are in one-to-one correspondence with unirreps of $\widetilde{G}_\phi$, which in turn define projective unirreps of G_ϕ. While have seen in Sec. F.2 that (projective) unirreps of G_ϕ also define projective unirreps of G, the construction does not necessarily go in the other direction, i.e., a projective unirrep of G does not necessarily define a unirrep of G_ϕ. This is because the association $g \mapsto U(g)$ is local, so it may not be possible to define a map $\phi : G \times G \to \mathbb{R}$ on the entire domain. Nevertheless, such a topological aspect is inconsequential at the level of the algebras, so a projective unirrep of G does define a self-adjoint irrep of $\mathfrak{g}_\varphi$, for some 2-cocycle $\varphi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$. This closes the loop, as depicted in the following diagram:



In conclusion, projective unirreps of a group G are in one-to-one correspondence with unirreps of the universal cover of the central extensions (by 2-cocycles ϕ) of the group, $\widetilde{G}_\phi$, and also in one-to-one correspondence with self-adjoint irreps of the central extensions (by 2-cocycles φ) of its Lie algebra, $\mathfrak{g}_\varphi$.

As an example, let us consider a rotation-invariant physical system in three Euclidean dimensions. The group of symmetry is $SO(3)$. It happens that its algebra, $\mathfrak{so}(3)$, does not admit (non-trivial) central extensions by 2-cocycles (in fact, this is true for any semisimple algebra). So the projective unirreps of $SO(3)$ must be in correspondence with true unirreps of its universal cover, $SU(2)$. Suppose one continuously rotate the system around some axis until it goes around by 2π. This curve starts and ends at $e \in SO(3)$, but it is not contractible. Viewing this process in $SU(2)$ the curve is open, as it starts at e and ends at e' (i.e., goes from the north to the south pole in $SU(2) \sim S^3$). Since e' is in the center, $U(e') = e^{i\alpha}$, and since $(e')^2 = e$, then α is 0 or π. That is, either the state returns to its original value (boson), $\Psi \mapsto \Psi$, or it flips sign (fermion), $\Psi \mapsto -\Psi$.

References

[1] R. Andrade e Silva, "Quantization of causal diamonds in (2+1)-dimensional gravity, Part I: Classical reduction," *arXiv:2308.11741*, 2023.

[2] R. Andrade e Silva, "Quantization of causal diamonds in (2+1)-dimensional gravity, Part II: Group-theoretic quantization," *arXiv:2310.03100*, 2023.

[3] R. Andrade e Silva and T. Jacobson, "Causal diamonds in (2+1)-dimensional quantum gravity," *Physical Review D*, vol. 107, no. 2, p. 024033, 2023.

[4] E. Witten, "2+1 dimensional gravity as an exactly soluble system," *Nuclear Physics B*, vol. 311, no. 1, pp. 46–78, 1988.

[5] C. J. Isham, "Canonical quantum gravity and the problem of time," in *Integrable systems, quantum groups, and quantum field theories*, pp. 157–287, Springer, 1993.

[6] E. Anderson, "Problem of time in quantum gravity," *Annalen der Physik*, vol. 524, no. 12, pp. 757–786, 2012.

[7] K. V. Kuchař, "Time and interpretations of quantum gravity," *International Journal of Modern Physics D*, vol. 20, no. supp01, pp. 3–86, 2011.

[8] V. Moncrief, "How solvable is (2+1)-dimensional Einstein gravity?," *Journal of Mathematical Physics*, vol. 31, no. 12, pp. 2978–2982, 1990.

[9] V. Moncrief, "Reduction of the Einstein equations in 2+1 dimensions to a Hamiltonian system over Teichmüller space," *Journal of Mathematical Physics*, vol. 30, no. 12, pp. 2907–2914, 1989.

[10] A. E. Fischer and V. Moncrief, "Hamiltonian reduction of Einstein's equations of general relativity," *Nuclear Physics B-Proceedings Supplements*, vol. 57, no. 1-3, pp. 142–161, 1997.

[11] J. E. Marsden and F. J. Tipler, "Maximal hypersurfaces and foliations of constant mean curvature in general relativity," *Physics Reports*, vol. 66, no. 3, pp. 109–139, 1980.

[12] A. E. Fischer and J. E. Marsden, "Linearization stability of nonlinear," *Differential Geometry, Part 2*, vol. 27, p. 219, 1975.

[13] D. Brill, *Proceedings of the First Marcel Grossmann Meeting on General Relativity.* North-Holland, (ed. R. Ruffini) 1977.

[14] R. Schoen and S.-T. Yau, "Incompressible minimal surfaces, three-dimensional manifolds with nonnegative scalar curvature, and the positive mass conjecture in general relativity," *Proceedings of the National Academy of Sciences*, vol. 75, no. 6, pp. 2567–2567, 1978.

[15] R. Schoen and S.-T. Yau, "Existence of incompressible minimal surfaces and the topology of three dimensional manifolds with non-negative scalar curvature," *Annals of Mathematics*, vol. 110, no. 1, pp. 127–142, 1979.

[16] Y. Choquet-Bruhat, "Einstein constraints on compact n-dimensional manifolds," *Classical and Quantum Gravity*, vol. 21, no. 3, p. S127, 2004.

[17] E. Witten, "Topology-changing amplitudes in 2+1 dimensional gravity," *Nuclear Physics B*, vol. 323, no. 1, pp. 113–140, 1989.

[18] A. Ashtekar, V. Husain, C. Rovelli, J. Samuel, and L. Smolin, "2+1 quantum gravity as a toy model for the 3+1 theory," *Classical and Quantum Gravity*, vol. 6, no. 10, p. L185, 1989.

[19] A. Hosoya and K. Nakao, "(2+1)-dimensional pure gravity for an arbitrary closed initial surface," *Classical and Quantum Gravity*, vol. 7, no. 2, p. 163, 1990.

[20] S. Carlip, *Quantum gravity in 2+1 dimensions.* Cambridge Monographs on Mathematical Physics, Cambridge University Press, 12 2003.

[21] S. Carlip, "Quantum gravity in 2+1 dimensions: the case of a closed universe," *Living Reviews in Relativity*, vol. 8, no. 1, pp. 1–63, 2005.

[22] P. Mondal, "Thurston boundary of the teichm\" uller space is the space of big bang singularities of 2+1 gravity," *arXiv preprint arXiv:2002.03551*, 2020.

[23] P. Kraus, R. Monten, and R. M. Myers, "3d gravity in a box," *SciPost Physics*, vol. 11, no. 3, p. 070, 2021.

[24] H. Adami, V. Hosseinzadeh, and M. Sheikh-Jabbari, "Sliding surface charges on AdS3," *Physics Letters B*, vol. 806, p. 135503, 2020.

[25] S. Ebert, E. Hijano, P. Kraus, R. Monten, and R. M. Myers, "Field theory of interacting boundary gravitons," *arXiv preprint arXiv:2201.01780*, 2022.

[26] J. D. Brown and M. Henneaux, "Central charges in the canonical realization of asymptotic symmetries: an example from three dimensional gravity," *Communications in Mathematical Physics*, vol. 104, no. 2, pp. 207–226, 1986.

[27] L. Freidel, J. Kowalski-Glikman, and L. Smolin, "2+1 gravity and doubly special relativity," *Physical Review D*, vol. 69, no. 4, p. 044001, 2004.

[28] S. Carlip, "Conformal field theory, (2+1)-dimensional gravity and the BTZ black hole," *Classical and Quantum Gravity*, vol. 22, no. 12, p. R85, 2005.

[29] E. Witten, "Three-dimensional gravity revisited," *arXiv preprint arXiv:0706.3359*, 2007.

[30] A. Maloney and E. Witten, "Quantum gravity partition functions in three dimensions," *Journal of High Energy Physics*, vol. 2010, no. 2, pp. 1–58, 2010.

[31] C. Scarinci and K. Krasnov, "The universal phase space of AdS_3 gravity," *Commun. Math. Phys.*, vol. 322, pp. 167–205, 2013.

[32] J. Kim and M. Porrati, "On a canonical quantization of 3D Anti de Sitter pure gravity," *Journal of High Energy Physics*, vol. 2015, no. 10, pp. 1–52, 2015.

[33] J. Cotler and K. Jensen, "A theory of reparameterizations for AdS3 gravity," *Journal of High Energy Physics*, vol. 2019, no. 2, pp. 1–73, 2019.

[34] V. Chandrasekaran and K. Prabhu, "Symmetries, charges and conservation laws at causal diamonds in general relativity," *Journal of High Energy Physics*, vol. 2019, no. 10, pp. 1–27, 2019.

[35] J. de Boer, F. M. Haehl, M. P. Heller, and R. C. Myers, "Entanglement, holography and causal diamonds," *Journal of High Energy Physics*, vol. 2016, no. 8, pp. 1–83, 2016.

[36] T. Banks, P. Draper, and S. Farkas, "Path integrals for causal diamonds and the covariant entropy principle," *Physical Review D*, vol. 103, no. 10, p. 106022, 2021.

[37] T. Jacobson and M. R. Visser, "Gravitational thermodynamics of causal diamonds in (a) ds," *SciPost Physics*, vol. 7, no. 6, p. 079, 2019.

[38] T. Jacobson and M. R. Visser, "Entropy of causal diamond ensembles," *arXiv preprint hep-th/2212.10608*, 2022.

[39] E. Witten, "A note on the canonical formalism for gravity," *arXiv preprint arXiv:2212.08270*, 2022.

[40] J. W. York Jr, "Role of conformal three-geometry in the dynamics of gravitation," *Physical review letters*, vol. 28, no. 16, p. 1082, 1972.

[41] C. J. Isham, "Topological and global aspects of quantum theory," in *Relativity, groups and topology. 2* (B. S. DeWitt and R. Stora, eds.), North-Holland Physics Pub., 1984.

[42] C. Isham, "Canonical groups and the quantization of general relativity," *Nuclear Physics B-Proceedings Supplements*, vol. 6, pp. 349–356, 1989.

[43] G. Barnich and G. Compere, "Classical central extension for asymptotic symmetries at null infinity in three spacetime dimensions," *Classical and Quantum Gravity*, vol. 24, no. 5, p. F15, 2007.

[44] G. W. Mackey, *Induced representations of groups and quantum mechanics.* New York, NY: Benjamin, 1968.

[45] B. Oblak, *BMS particles in three dimensions.* Springer, 2017.

[46] G. W. Mackey, "Infinite-dimensional group representations," *Bulletin of the American Mathematical Society*, vol. 69, no. 5, pp. 628–686, 1963.

[47] R. Andrade e Silva and T. Jacobson, "Particle on the sphere: group-theoretic quantization in the presence of a magnetic monopole," *Journal of Physics A: Mathematical and Theoretical*, vol. 54, no. 23, p. 235303, 2021.

[48] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Princeton university press, 1992.

[49] A. Ashtekar and G. T. Horowitz, "On the canonical approach to quantum gravity," *Physical Review D*, vol. 26, no. 12, p. 3342, 1982.

[50] R. Arnowitt, S. Deser, and C. W. Misner, "Republication of: The dynamics of general relativity," *General Relativity and Gravitation*, vol. 40, no. 9, pp. 1997–2027, 2008.

[51] A. E. Fischer and V. Moncrief, "The reduced einstein equations and the conformal volume collapse of 3-manifolds," *Classical and Quantum Gravity*, vol. 18, no. 21, p. 4493, 2001.

[52] A. J. Speranza, "Local phase space and edge modes for diffeomorphism-invariant theories," *Journal of High Energy Physics*, vol. 2018, no. 2, 2018.

[53] A. J. Speranza, "Ambiguity resolution for integrable gravitational charges," *Journal of High Energy Physics*, vol. 2022, no. 7, pp. 1–23, 2022.

[54] T. Jacobson, "Black hole entropy and induced gravity," *arXiv preprint gr-qc/9404039*, 1994.

[55] D. Harlow and J.-q. Wu, "Covariant phase space with boundaries," *Journal of High Energy Physics*, vol. 2020, no. 10, pp. 1–52, 2020.

[56] V. Chandrasekaran and A. J. Speranza, "Anomalies in gravitational charge algebras of null boundaries and black hole entropy," *Journal of High Energy Physics*, vol. 2021, no. 1, pp. 1–56, 2021.

[57] J. Kirklin, "Unambiguous phase spaces for subregions," *Journal of High Energy Physics*, vol. 2019, no. 3, pp. 1–24, 2019.

[58] A. Lichnerowicz, *L'intégration des équations de la gravitation relativiste et le problème des n-corps.* Gauthier-Villars, 1944.

[59] A. Lichnerowicz, "Sur les équations relativistes de la gravitation," *Bulletin de la Société Mathématique de France*, vol. 80, pp. 237–251, 1952.

[60] Y. Choquet-Bruhat, *General relativity and the Einstein equations.* OUP Oxford, 2008.

[61] R. Bartnik, "Regularity of variational maximal surfaces," *Acta Mathematica*, vol. 161, pp. 145–181, 1988.

[62] R. Bartnik, "Remarks on cosmological spacetimes and constant mean curvature surfaces," *Communications in mathematical physics*, vol. 117, no. 4, pp. 615–624, 1988.

[63] D. Brill and F. Flaherty, "Isolated maximal surfaces in spacetime," *Communications in Mathematical Physics*, vol. 50, no. 2, pp. 157–165, 1976.

[64] C. Gerhardt, "H-surfaces in Lorentzian manifolds," *Communications in mathematical physics*, vol. 89, no. 4, pp. 523–553, 1983.

[65] D. M. Eardley and L. Smarr, "Time functions in numerical relativity: marginally bound dust collapse," *Physical Review D*, vol. 19, no. 8, p. 2239, 1979.

[66] H. Friedrich, I. Racz, and R. M. Wald, "On the rigidity theorem for spacetimes with a stationary event horizon or a compact cauchy horizon," *Communications in mathematical physics*, vol. 204, pp. 691–707, 1999.

[67] N. O'Murchadha and J. W. York Jr, "Existence and uniqueness of solutions of the hamiltonian constraint of general relativity on compact manifolds," *Journal of Mathematical Physics*, vol. 14, no. 11, pp. 1551–1557, 1973.

[68] L. C. Evans, *Partial differential equations*, vol. 19. American Mathematical Society, 2022.

[69] C. J. Isham, *Modern differential geometry for physicists*, vol. 61. World Scientific Publishing Company, 1999.

[70] S. Helgason, *Differential geometry and symmetric spaces*, vol. 341. American Mathematical Soc., 2001.

[71] J. Bak, D. J. Newman, and D. J. Newman, *Complex analysis*, vol. 8. Springer, 2010.

[72] J. Liouville, "Sur l'équation aux différences partielles $\frac{d^2 \log \lambda}{dudv} \pm \frac{\lambda}{2a^2} = 0$," *Journal de mathématiques pures et appliquées*, vol. 18, pp. 71–72, 1853.

[73] N. Seiberg, "Notes on quantum liouville theory and quantum gravity," *Progress of Theoretical Physics Supplement*, vol. 102, pp. 319–349, 1990.

[74] J. Balog, L. Fehér, and L. Palla, "Coadjoint orbits of the Virasoro algebra and the global Liouville equation," *International Journal of Modern Physics A*, vol. 13, no. 02, pp. 315–362, 1998.

[75] K. Krasnov, "Three-dimensional gravity, point particles and liouville theory," *Classical and Quantum Gravity*, vol. 18, no. 7, p. 1291, 2001.

[76] Y. Nakayama, "Liouville field theory: A decade after the revolution," *International Journal of Modern Physics A*, vol. 19, no. 17n18, pp. 2771–2930, 2004.

[77] L. D. Faddeev, "Zero modes for the quantum liouville model," *arXiv preprint arXiv:1404.1713*, 2014.

[78] H. Erbin, "Notes on 2d quantum gravity and liouville theory," 2015.

[79] S. Li, N. Toumbas, and J. Troost, "Liouville quantum gravity," *Nuclear Physics B*, vol. 952, p. 114913, 2020.

[80] S. T. Ali and M. Engliš, "Quantization methods: a guide for physicists and analysts," *Reviews in Mathematical Physics*, vol. 17, no. 04, pp. 391–490, 2005.

[81] M. J. Gotay, "Obstructions to quantization," in *Mechanics: from theory to computation*, pp. 171–216, Springer, 2000.

[82] M. J. Gotay, "On a full quantization of the torus," in *Quantization, coherent states, and complex structures*, pp. 55–62, Springer, 1995.

[83] M. J. Gotay and J. Grabowski, "On quantizing nilpotent and solvable basic algebras," *Canadian Mathematical Bulletin*, vol. 44, no. 2, pp. 140–149, 2001.

[84] B. C. Hall, *Quantum theory for mathematicians.* Springer, 2013.

[85] N. M. J. Woodhouse, *Geometric quantization.* Oxford university press, 1992.

[86] R. Dudley, J. Feldman, B. Kostant, R. Langlands, E. Stein, and B. Kostant, *Quantization and unitary representations.* Springer, 1970.

[87] B. Kostant, "On the definition of quantization," *Géométrie Symplectique et Physique Mathématique*, vol. 237, 1974.

[88] J.-M. Souriau, "Modèle de particule à spin dans le champ électromagnétique et gravitationnel," in *Annales de l'institut Henri Poincaré. Section A, Physique Théorique*, vol. 20, pp. 315–364, 1974.

[89] A. A. Kirillov, "Geometric quantization," in *Dynamical Systems IV: Symplectic Geometry and its Applications*, pp. 139–176, Springer, 2001.

[90] D. Sternheimer, "Deformation quantization: Twenty years after," in *AIP Conference Proceedings*, vol. 453, pp. 107–145, American Institute of Physics, 1998.

[91] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, "Deformation theory and quantization. i. deformations of symplectic structures," *Annals of Physics*, vol. 111, no. 1, pp. 61–110, 1978.

[92] C. J. Isham and A. C. Kakas, "A group theoretical approach to the canonical quantisation of gravity. i. construction of the canonical group," *Classical and Quantum Gravity*, vol. 1, no. 6, p. 621, 1984.

[93] C. Isham and A. C. Kakas, "A group theoretical approach to the canonical quantisation of gravity. ii. unitary representations of the canonical group," *Classical and Quantum Gravity*, vol. 1, no. 6, p. 633, 1984.

[94] E. Witten, "Coadjoint orbits of the Virasoro group," *Communications in Mathematical Physics*, vol. 114, no. 1, pp. 1–53, 1988.

[95] V. F. Lazutkin and T. F. Pankratova, "Normal forms and versal deformations for Hill's equation," *Functional Analysis and its applications*, vol. 9, no. 4, pp. 306–311, 1975.

[96] G. Segal, "Unitary representations of some infinite dimensional groups," *Communications in Mathematical Physics*, vol. 80, no. 3, pp. 301–342, 1981.

[97] A. Alekseev and S. Shatashvili, "Path integral quantization of the coadjoint orbits of the virasoro group and 2-d gravity," *Nuclear Physics B*, vol. 323, no. 3, pp. 719–733, 1989.

[98] L. Guieu and C. Roger, *L'Algèbre et le Groupe de Virasoro.* Les Publications CRM (Centre de Recherches Mathématiques de Montréal), 2007.

[99] B. A. Khesin and R. Wendt, *The geometry of infinite-dimensional groups*, vol. 51. Springer, 2009.

[100] V. Ovsienko and S. Tabachnikov, *Projective differential geometry old and new: from the Schwarzian derivative to the cohomology of diffeomorphism groups*, vol. 165. Cambridge University Press, 2004.

[101] A. V. Kosyak, *Regular, quasi-regular and induced representations of infinite-dimensional groups.* 2018.

[102] W. Donnelly, L. Freidel, S. F. Moosavian, and A. J. Speranza, "Gravitational edge modes, coadjoint orbits, and hydrodynamics," *Journal of High Energy Physics*, vol. 2021, no. 9, pp. 1–63, 2021.

[103] A. Castro, S. Detournay, N. Iqbal, and E. Perlmutter, "Holographic entanglement entropy and gravitational anomalies," *Journal of High Energy Physics*, vol. 2014, no. 7, pp. 1–52, 2014.

[104] B. Oblak, "Characters of the BMS group in three dimensions," *Communications in Mathematical Physics*, vol. 340, no. 1, pp. 413–432, 2015.

[105] G. Barnich and B. Oblak, "Notes on the BMS group in three dimensions: I. Induced representations," *Journal of High Energy Physics*, vol. 2014, no. 6, pp. 1–27, 2014.

[106] G. Barnich and B. Oblak, "Notes on the BMS group in three dimensions: II. Coadjoint representation," *Journal of High Energy Physics*, vol. 2015, no. 3, pp. 1–18, 2015.

[107] A. Campoleoni, H. A. Gonzalez, B. Oblak, and M. Riegler, "BMS modules in three dimensions," *International Journal of Modern Physics A*, vol. 31, no. 12, p. 1650068, 2016.

[108] V. Bargmann, "Irreducible unitary representations of the lorentz group," *Annals of Mathematics*, pp. 568–640, 1947.

[109] I. M. Gel'fand and M. A. Naimark, "Unitary representations of the lorentz group," *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 11, no. 5, pp. 411–504, 1947.

[110] Harish-Chandra, "Plancherel formula for the 2× 2 real unimodular group," *Proceedings of the National Academy of Sciences*, vol. 38, no. 4, pp. 337–342, 1952.

[111] A. Kitaev, "Notes on sl (2, r) representations," *arXiv preprint arXiv:1711.08169*, 2017.

[112] J. Hilgert and K.-H. Neeb, *Structure and geometry of Lie groups.* Springer Science & Business Media, 2011.

[113] R. de la Madrid Modino, *Quantum mechanics in rigged Hilbert space language*. PhD thesis, PhD thesis. Universidat de Valladolid, 2001. url: http://galaxy. cs. lamar . . . , 2001.

[114] I. Gelfand and N. YA, "Vilenkin, generalized functions, iv," 1964.

[115] L. Cornalba and M. S. Costa, "A new cosmological scenario in string theory," *Physical Review D*, vol. 66, no. 6, p. 066001, 2002.

[116] L. Cornalba and M. S. Costa, "Time-dependent orbifolds and string cosmology," *arXiv preprint hep-th/0310099*, 2003.

[117] G. Barnich and B. Oblak, "Holographic positive energy theorems in three-dimensional gravity," *Classical and Quantum Gravity*, vol. 31, no. 15, p. 152001, 2014.

[118] R. Schwartz, "A projectively natural flow for circle diffeomorphisms," *Inventiones mathematicae*, vol. 110, no. 1, pp. 627–647, 1992.

[119] S. Kulkarni, M. Nair, and G. Ramesh, "Some properties of unbounded operators with closed range," *Proceedings Mathematical Sciences*, vol. 118, pp. 613–625, 2008.

[120] V. I. Bogachev and M. A. S. Ruas, *Measure theory*, vol. 1. Springer, 2007.

[121] S. Carlip and S. J. Carlip, *Quantum gravity in 2+1 dimensions*, vol. 50. Cambridge University Press, 2003.

[122] S.-S. Chern and J. Simons, "Characteristic forms and geometric invariants," *Annals of Mathematics*, vol. 99, no. 1, pp. 48–69, 1974.

[123] T. Jacobson, "Entanglement equilibrium and the Einstein equation," *Physical review letters*, vol. 116, no. 20, p. 201101, 2016.

[124] T. Jacobson and M. Visser, "Spacetime equilibrium at negative temperature and the attraction of gravity," *International Journal of Modern Physics D*, vol. 28, no. 14, p. 1944016, 2019.

[125] T. Jacobson and M. R. Visser, "Entropy of causal diamond ensembles," *SciPost Physics*, vol. 15, no. 1, p. 023, 2023.

[126] T. Jacobson and M. R. Visser, "Partition function for a volume of space," *Physical Review Letters*, vol. 130, no. 22, p. 221501, 2023.

[127] T. Banks and P. Draper, "Comments on the entanglement spectrum of de sitter space," *Journal of High Energy Physics*, vol. 2023, no. 1, pp. 1–10, 2023.

[128] R. Bousso, "A covariant entropy conjecture," *Journal of High Energy Physics*, vol. 1999, no. 07, p. 004, 1999.

[129] R. Bousso, "Holography in general space-times," *Journal of High Energy Physics*, vol. 1999, no. 06, p. 028, 1999.

[130] W. Fischler and L. Susskind, "Holography and cosmology," *arXiv preprint hep-th/9806039*, 1998.

[131] P. Martinetti and C. Rovelli, "Diamond's temperature: Unruh effect for bounded trajectories and thermal time hypothesis," *Classical and Quantum Gravity*, vol. 20, no. 22, p. 4919, 2003.

[132] C. Isham, "Topos methods in the foundations of physics," *Deep beauty*, pp. 187–206, 2011.

[133] A. Döring and C. J. Isham, "A topos foundation for theories of physics: I. formal languages for physics," *Journal of Mathematical Physics*, vol. 49, no. 5, 2008.

[134] A. Döring and C. Isham, ""what is a thing?": Topos theory in the foundations of physics," in *New structures for physics*, pp. 753–937, Springer, 2010.

[135] C. P. Burgess, "Quantum gravity in everyday life: General relativity as an effective field theory," *Living Reviews in Relativity*, vol. 7, pp. 1–56, 2004.

[136] D. Baumann, "Tasi lectures on inflation," *arXiv preprint arXiv:0907.5424*, 2009.

[137] D. Carney, P. C. Stamp, and J. M. Taylor, "Tabletop experiments for quantum gravity: a user's manual," *Classical and Quantum Gravity*, vol. 36, no. 3, p. 034001, 2019.

[138] A. M. Vershik, I. M. Gel'fand, and M. I. Graev, “Representations of the group of diffeomorphisms,” *Russian Mathematical Surveys*, vol. 30, no. 6, p. 1, 1975.

[139] R. Andrade e Silva, “Unitarity of time evolution for a particle on the half-line,” *In preparation*, 2024.

[140] A. Vilenkin, “Approaches to quantum cosmology,” *Physical Review D*, vol. 50, p. 2581–2594, Aug. 1994.

[141] N. Pinto-Neto and J. Fabris, “Quantum cosmology from the de broglie–bohm perspective,” *Classical and Quantum Gravity*, vol. 30, no. 14, p. 143001, 2013.

[142] N. Pinto-Neto, F. Falciano, R. Pereira, and E. S. Santini, “Wheeler-dewitt quantization can solve the singularity problem,” *Physical Review D*, vol. 86, no. 6, p. 063504, 2012.

[143] S. P. Kim, “Quantum potential and cosmological singularities,” *Physics Letters A*, vol. 236, p. 11–15, Dec. 1997.

[144] D. Harlow and D. Jafferis, “The factorization problem in jackiw-teitelboim gravity,” *Journal of High Energy Physics*, vol. 2020, no. 2, pp. 1–32, 2020.

[145] J. Klauder and E. Aslaksen, “Elementary model for quantum gravity,” *Phys. Rev. D*, vol. 2, pp. 272–276, 1970.

[146] M. Pilati, “Strong-coupling quantum gravity. I. Solution in a particular gauge,” *Physical Review D*, vol. 26, no. 10, p. 2645, 1982.

[147] M. Pilati, "Strong-coupling quantum gravity. II. Solution without gauge fixing," *Physical Review D*, vol. 28, no. 4, p. 729, 1983.

[148] J. R. Klauder, "Noncanonical quantization of gravity. I. Foundations of affine quantum gravity," *Journal of Mathematical Physics*, vol. 40, no. 11, pp. 5860–5882, 1999.

[149] J. R. Klauder, "The affine quantum gravity programme," *Classical and Quantum Gravity*, vol. 19, no. 4, p. 817, 2002.

[150] R. Andrade e Silva, "A non-perturbative setup for quantizing the non-linear sigma model," *In preparation*, 2024.

[151] W. Donnelly and S. B. Giddings, "Observables, gravitational dressing, and obstructions to locality and subsystems," *Physical Review D*, vol. 94, no. 10, p. 104038, 2016.

[152] W. Donnelly and S. B. Giddings, "Diffeomorphism-invariant observables and their nonlocal algebra," *Physical Review D*, vol. 93, no. 2, p. 024030, 2016.

[153] W. Donnelly and S. B. Giddings, "How is quantum information localized in gravity?," *Physical Review D*, vol. 96, no. 8, p. 086013, 2017.

[154] J. S. Cotler, G. R. Penington, and D. H. Ranard, "Locality from the spectrum," *Communications in Mathematical Physics*, vol. 368, pp. 1267–1296, 2019.

[155] G. C. Hegerfeldt, "Remark on causality and particle localization," *Physical Review D*, vol. 10, no. 10, p. 3320, 1974.

[156] W. Donnelly and L. Freidel, "Local subsystems in gauge theory and gravity," *Journal of High Energy Physics*, vol. 2016, no. 9, pp. 1–45, 2016.

[157] W. Donnelly, L. Freidel, S. F. Moosavian, and A. J. Speranza, "Matrix quantization of gravitational edge modes," *Journal of High Energy Physics*, vol. 2023, no. 5, pp. 1–95, 2023.

[158] M. Geiller, "Edge modes and corner ambiguities in 3d chern–simons theory and gravity," *Nuclear Physics B*, vol. 924, pp. 312–365, 2017.

[159] M. Geiller, "Lorentz-diffeomorphism edge modes in 3d gravity," *Journal of High Energy Physics*, vol. 2018, no. 2, 2018.

[160] L. Freidel, M. Geiller, and D. Pranzetti, "Edge modes of gravity. part i. corner potentials and charges," *Journal of High Energy Physics*, vol. 2020, no. 11, pp. 1–52, 2020.

[161] L. Freidel, M. Geiller, and D. Pranzetti, "Edge modes of gravity. part ii. corner metric and lorentz charges," *Journal of High Energy Physics*, vol. 2020, no. 11, pp. 1–64, 2020.

[162] L. Freidel, M. Geiller, and D. Pranzetti, "Edge modes of gravity. part iii. corner simplicity constraints," *Journal of High Energy Physics*, vol. 2021, no. 1, pp. 1–64, 2021.

[163] A. Strominger and C. Vafa, “Microscopic origin of the bekenstein-hawking entropy,” *Physics Letters B*, vol. 379, no. 1-4, pp. 99–104, 1996.

[164] P. Pulakkat, “On the charge algebra of causal diamonds in three dimensional gravity,” *arXiv preprint arXiv:2404.03014*, 2024.

[165] R. M. Wald, “Black hole entropy is the noether charge,” *Physical Review D*, vol. 48, no. 8, p. R3427, 1993.

[166] H.-J. Matschull and M. Welling, “Quantum mechanics of a point particle in (2+1)-dimensional gravity,” *Classical and Quantum Gravity*, vol. 15, no. 10, p. 2981, 1998.

[167] H.-J. Matschull, “Black hole creation in 2+1 dimensions,” *Classical and Quantum Gravity*, vol. 16, no. 3, p. 1069, 1999.

[168] S. Carlip, “Exact quantum scattering in 2+ 1 dimensional gravity,” *Nuclear Physics B*, vol. 324, no. 1, pp. 106–122, 1989.

[169] J. Raeymaekers, “Quantization of conical spaces in 3d gravity,” *Journal of High Energy Physics*, vol. 2015, no. 3, pp. 1–28, 2015.

[170] M. Banados, C. Teitelboim, and J. Zanelli, “Black hole in three-dimensional spacetime,” *Physical review letters*, vol. 69, no. 13, p. 1849, 1992.

[171] N. Benjamin, S. Collier, and A. Maloney, “Pure gravity and conical defects,” *Journal of High Energy Physics*, vol. 2020, no. 9, pp. 1–31, 2020.

[172] T. Frankel, *The geometry of physics: an introduction.* Cambridge university press, 2011.

[173] G. W. Mackey, "Induced representations of locally compact groups i," *Annals of Mathematics*, pp. 101–139, 1952.

[174] R. Raczka and A. O. Barut, *Theory of group representations and applications.* World Scientific Publishing Company, 1986.

[175] N. Wiener, "Differential-space," *Journal of Mathematics and Physics*, vol. 2, no. 1-4, pp. 131–174, 1923.

[176] P. Kuzmin, "On circle diffeomorphisms with discontinuous derivatives and quasi-invariance subgroups of malliavin–shavgulidze measures," *Journal of mathematical analysis and applications*, vol. 330, no. 1, pp. 744–750, 2007.

[177] E. Shavgulidze, "Distributions on infinite-dimensional spaces and second quantization in string theories, ii," in *V International Vilnius Conference on Probability Theory and Math. Statistics, Abstracts of Comm., Vilnius*, pp. 359–360, 1989.

[178] A. Kosyak, "Irreducible regular gaussian representations of the groups of the interval and circle diffeomorphisms," *Journal of Functional Analysis*, vol. 125, no. 2, pp. 493–547, 1994.

[179] G. Mess, "Lorentz spacetimes of constant curvature," *arXiv preprint arXiv:0706.1570*, 2007.

[180] K. Krasnov and J.-M. Schlenker, "Minimal surfaces and particles in 3-manifolds," *Geometriae Dedicata*, vol. 126, pp. 187–254, 2007.

[181] F. Bonsante and J.-M. Schlenker, "Maximal surfaces and the universal teichmüller space," *Inventiones mathematicae*, vol. 182, no. 2, pp. 279–333, 2010.

[182] C. Mackey, "Induced representations," *Representation Theory of Lie Groups*, p. 20, 1965.

[183] G. W. Mackey, "Unitary group representations in physics, probability, and number theory," *(No Title)*, 1978.

[184] G. W. Mackey, "Induced representations of locally compact groups ii. the frobenius reciprocity theorem," *Annals of Mathematics*, pp. 193–221, 1953.

[185] N. Bourbaki, *Elements of Mathematics: Integration*. Addison-Wesley, 1975.

[186] G. B. Folland, *A course in abstract harmonic analysis*, vol. 29. CRC press, 2016.

[187] V. Bargmann, "On unitary ray representations of continuous groups," *Annals of Mathematics*, pp. 1–46, 1954.

[188] ACuriousMind, "Why exactly do sometimes universal covers, and sometimes central extensions feature in the application of a symmetry group to quantum physics?." Physics Stack Exchange. https://physics.stackexchange.com/q/203945 (version: 2020-06-11).

[189] E. P. Wigner, "Gruppentheorie und ihre anwendung auf die quantenmechanik der atomspektren," 1931.

[190] S. Weinberg, *The quantum theory of fields: Volume 1, foundations*. Cambridge university press, 2005.