

Gravitational Hilbert spaces: invariant and co-invariant states, inner products, gauge-fixing, and BRST

Jesse Held,^a Henry Maxfield^b

^a*Department of Physics, University of California Santa Barbara, USA*

^b*Leinweber Institute for Theoretical Physics at Stanford, 382 Via Pueblo, Stanford, CA 94305, USA*

E-mail: jheld@ucsb.edu, henrym@stanford.edu

ABSTRACT: Hilbert spaces in theories of gravity are notoriously subtle due to the Hamiltonian constraints, particularly regarding the inner product. To demystify this subject, we review and extend a collection of ideas in canonical gravity, and connect to the sum-over-histories approach by clarifying the Hilbert space interpretation of various gravitational path integrals. We use one-dimensional (or mini-superspace) models as the simplest context to exemplify the conceptual ideas. We emphasise that a physical Hilbert space can be defined either by requiring states to be annihilated by constraint operators (e.g., the Wheeler-DeWitt equation) or by equivalence relations between wavefunctions, and explain that these two approaches are related by an inner product. We advocate that the group averaging procedure constructs the correct physical inner product. The Klein-Gordon inner product is not positive-definite, which we explain as arising from a bad gauge choice; nonetheless, it agrees with group averaging when such a problem is absent. These concepts are all embedded in the BRST/BFV formalism, which provides a systematic way to construct these and other physically equivalent inner products (e.g., from maximal-volume gauge and Gaussian averaged gauges). Finally we discuss the application of these ideas in the semi-classical approximation, including non-perturbative gravitational effects.

Contents

1	Introduction and detailed summary	1
1.1	States in quantum gravity: not (only) Wheeler-DeWitt	3
1.2	Co-invariants, gauge-fixing, and null states	5
1.3	Gravitational inner products	6
1.4	Group averaging	7
1.5	Klein-Gordon and more inner products	8
1.6	Gravitational path integrals	10
1.7	Inner products in the BRST formalism	12
1.8	Perturbative and non-perturbative physics in the semi-classical limit	14
1.9	Paper outline	15
2	One-dimensional (or mini-superspace) gravity	16
2.1	Coupling quantum mechanics to one-dimensional gravity	16
2.2	Examples: sigma models	17
2.3	One dimensional gravity vs worldline QFT	18
3	States: invariants and co-invariants	18
3.1	States and the Hamiltonian constraint	18
3.2	Technical comments	19
3.3	Invariants and co-invariants in BRST quantisation	20
3.4	Natural invariant and co-invariant states	22
3.5	States in the path integral formalism	24
3.6	Co-invariants and gauge-fixing	25
4	Gravitational inner products	27
4.1	Inner products and equivalences between invariants and co-invariants	27
4.2	Group averaging inner product	29
4.3	The Klein-Gordon form	30
4.4	Path integrals, inner products and gauge fixing	33
4.5	Example: a one-dimensional target space	37
4.6	The QFT inner product	40
5	Operators on the physical Hilbert space	42
5.1	Dirac observables	43
5.2	Observables in BRST	44
6	Inner products from the BFV-BRST formalism	45
6.1	Sketch of the main ideas	45
6.2	The non-minimal formalism and gauge-fixing	47
6.3	Non-minimal BRST	48
6.4	Some examples	50

6.5	Mixing invariants and co-invariants	55
7	Semiclassical limits	57
7.1	General comments	57
7.2	One-dimensional target examples	63
8	Discussion	72
8.1	Additional considerations in higher dimensions	72
8.2	Singularities: constraints are not essentially self-adjoint	74
8.3	Null states	75
8.4	Topology change	76
A	Proof of Klein-Gordon causal propagator identity	77
B	Static target spaces	78
C	More on BRST calculations	80
D	Semi-classical calculations	82
D.1	Perturbative solution of WDW and inner product	82
D.2	Examples	83
E	String worldsheet analogy	84
F	Maxwell theory	86
F.1	Wavefunctional representation	87
F.2	Oscillator representation and Gupta-Bleuler	89

1 Introduction and detailed summary

Among numerous approaches to quantum theories of gravity, the oldest and most straightforward is a direct quantisation of a dynamical spacetime metric and matter fields. In modern terms this is best understood as an effective quantum theory, valid for sufficiently low energy scales and weakly curved spacetimes. It has recently become apparent that this effective theory is far more powerful than previously envisaged, yielding several surprising results on quantum aspects of black holes. Highlights include calculations of the entropy of Hawking radiation [1, 2], statistic of energy eigenvalues [3], and state counting of BPS black holes [4]. In many cases, these results rely on non-perturbative effects involving fluctuations in the topology of spacetime.

These results have largely been understood in the formalism of the gravitational path integral. While this has proven a very powerful tool, it has several drawbacks relative to the canonical formalism. A Hilbert space interpretation in bulk variables is not manifest, since the usual way to obtain intermediate states by cutting the path integral on a Cauchy

surface is not straightforward due to diffeomorphism invariance. Relatedly, a local (or approximately local) Lorentzian description of the dynamic is obscured, particularly when using a saddle-point approximation dominated by spacetimes with Euclidean or complex metrics. These features make it challenging to apply recent results to questions inherently about the bulk dynamics (rather than quantities defined asymptotically), most notably involving the black hole interior and quantum cosmology.

This motivates us to recast recent developments in the canonical formalism, and in a language which makes the bulk physics manifest. In particular, in the ‘semi-classical’ $G_N \rightarrow 0$ limit we aim for the closest possible approach to a standard Hilbert space description of local quantum field theory (LQFT). Explicitly, our aim is to formulate a gravitational Hilbert space satisfying the following criteria:

1. To leading order in the $G_N \rightarrow 0$ limit, the Hilbert space reduces to that of LQFT (including free gravitons) on a classical background spacetime.¹ In particular, this means we can define ‘Schrödinger picture’ states defined by wavefunctions of field configurations on a specific Cauchy surface.
2. Perturbative corrections can (at least in principle) be calculated order-by-order in G_N perturbation theory (up to the usual ambiguities from renormalisation of coupling constants).
3. Contributions from non-perturbative gravitational effects (e.g., saddle point geometries which are not close to the original background) can be incorporated.
4. The Hilbert space is equipped with a positive-definite inner product.

Furthermore, we would like to be able to construct gauge-invariant bulk observables acting on this Hilbert space, which reduce to local QFT operators in the $G_N \rightarrow 0$ limit (but require ‘gravitational dressing’ once gravity is present). We emphasise that we only attempt to describe physics in the regime of validity of gravitational effective field theory, so in this way our aims are more modest than much of the older literature on canonical gravity.

As already noted, the main challenge to such a description comes from diffeomorphism invariance and the associated constraints. For example, the common description of the Hilbert space as solutions to the Wheeler-DeWitt (WDW) equation does not (by itself) achieve our aims, because this description of states is not local in time (failing point (1) above). Point (4) emphasises that a Hilbert space requires not only a space of states, but also a positive-definite inner product; once again this is not straightforward because of the constraints (and is not satisfied by the commonly used Klein-Gordon (KG) inner product). This motivates us to explore alternatives which overcome these issues, and we will subsequently explain the relation to the better-known WDW/KG constructions.

The canonical methods we describe should not be viewed as a completely separate alternative to the path integral, but as a complementary language. In particular, the

¹If the background in question possesses isometries, then we should recover only the sector of the QFT Hilbert space invariant under the symmetries. We comment on this further in the discussion section 8.

canonical language will be useful to provide robust Hilbert space interpretations of various path integrals, especially those with boundary conditions at finite time (rather than asymptotically). It will also be helpful for identifying the correct measure associated with such path integrals, including path integrals which compute matrix elements of gauge-fixed bulk observables.

A specific goal of this program is to shed light on the idea of ‘null states’, which has become prominent in recent developments. This refers to a non-zero wavefunction of gravitational variables which (due to non-perturbative effects) has zero norm, leading to linear relations between apparently distinct states. From the perspective of the path integral, the existence of such states (but none with negative norm) appears to require a highly-tuned conspiracy, which demands some explanation. We advocate an explanation in terms of gauge symmetry, specifically non-perturbative diffeomorphism invariance (or perhaps a mild extension thereof). Null states can then result as the difference between two states related by a gauge transformation. One of our central motivations is to provide a language in which this idea can be made as precise and concrete as possible.

In this paper we describe a strategy for achieving these goals. We draw heavily on previous ideas in the literature on canonical quantum gravity, many of which are not as well-known as they deserve. As such, much of the paper will be dedicated to a review of this older work, emphasising the conceptual points which are most relevant to our aims, though in many places we extend these ideas and illustrate them with novel examples. We sketch the main ideas in the remainder of the introduction, while leaving details to the main text. To provide the simplest possible illustration and avoid excessive technical distractions, in the bulk of the paper we mostly flesh out the details in theories of one-dimensional (or mini-superspace) theories of gravity, which is sufficient to illustrate the salient points. Nonetheless, we concentrate on ideas which readily generalise to more interesting theories in higher dimensions (both closed universes and those with spatial asymptotic boundaries such as AdS).

The reader may find a specific solvable example useful: many of the main ideas in this paper are implemented and illustrated in detail for de Sitter JT gravity in [5].

1.1 States in quantum gravity: not (only) Wheeler-DeWitt

Gravity is exceptional among physical theories in that it does not have a conventional description of dynamics, due to the fact that time evolution is a gauge transformation (except at boundaries of spacetime). The Hamiltonian density \mathcal{H} does not generate time translations relating different states and observables, but rather imposes the Hamiltonian constraints $\mathcal{H} = 0$.² This raises both conceptual and technical issues, particularly if we wish to describe non-perturbative gravitational effects. These include defining the space of physical states, the inner product on those states, and observables which describe (approximately) local physics.

Starting with an unconstrained quantum theory with a Hilbert space \mathcal{H}_0 and constraint operators \mathcal{H} , we would like to construct a physical Hilbert space of the constrained theory

²Diffeomorphisms of a given Cauchy surface also give rise to momentum constraints $\mathcal{P} = 0$, though these are less challenging (both conceptually and technically) and we do not emphasise their role.

with $\mathcal{H} = 0$ (i.e., we quantise first and subsequently impose constraints).³ For gravity, \mathcal{H}_0 is the space of wavefunctionals $\psi[\gamma_{ij}(x)]$ of spatial metrics (as well as spatial configurations of matter fields). We first discuss the space of states, addressing the inner product later. In particular, we will not worry about normalisability of wavefunctions (and associated technical issues like completions to form a Hilbert space) since the relevant norm follows later from our choice of inner product.

The oldest and perhaps most well-known approach goes back to Dirac [6], and in the context of gravity the foundational work of DeWitt [7]. In this Dirac-DeWitt approach, the constraints are imposed as equations restricting the physical states, namely the Wheeler-DeWitt equation:

$$\mathcal{H}|\psi\rangle = 0 \quad (\text{invariant states } |\psi\rangle). \quad (1.1)$$

We call such states *invariants* since this condition expresses invariance of the wavefunction under the infinitesimal gauge transformations generated by the constraints, in this case changes of coordinates that change the initial $t = 0$ Cauchy surface. We use the notation $|\psi\rangle \in \mathcal{H}_{\text{inv}}$ for an invariant state to distinguish from arbitrary states in \mathcal{H}_0 .

However, solving the Wheeler-DeWitt equation is not the only way to construct a space of states. An alternative is to put no restriction on the physical wavefunctions, but instead to declare equivalences between wavefunctions related by action of the constraints:

$$|\psi\rangle \sim |\psi\rangle + \mathcal{H}|\zeta\rangle \quad (\text{co-invariant states } |\psi\rangle). \quad (1.2)$$

In this equation, \mathcal{H} can represent any linear combination of constraints, and $|\zeta\rangle$ an arbitrary state of the unconstrained theory. The physical states are cosets, equivalence classes of wavefunctions related by gauge transformations. We denote such an equivalence class as $|\psi\rangle\rangle \in \mathcal{H}_{\text{co}}$ (often using $|\psi\rangle\rangle$ to mean the equivalence class containing a state $|\psi\rangle \in \mathcal{H}_0$), and call these states *co-invariants* (following terminology in [8]). This name reflects the duality (in the sense of linear algebra) between the co-invariants and invariants: the inner product on the unconstrained Hilbert space induces a well-defined sesquilinear pairing between invariants and co-invariants (though not an inner product on either space).

$$\mathcal{H}_{\text{inv}} \text{ \& } \mathcal{H}_{\text{co}} \text{ are dual: } \llbracket \cdot | \cdot \rrbracket \text{ is a non-degenerate sesquilinear form } \overline{\mathcal{H}_{\text{inv}}} \otimes \mathcal{H}_{\text{co}} \rightarrow \mathbb{C}. \quad (1.3)$$

Standard approaches in other contexts (such as gauge theories and the string worldsheet) typically use a mixture of these two prescriptions for different constraints. For example, Gupta-Bleuler quantisation (or old covariant quantisation of the string) imposes invariance under positive frequency constraints and equivalence (co-invariance) under negative frequency constraints. BRST quantisation encompasses both depending on the state of ghosts. We expect such a mixture to be most practical for many applications to realistic theories of gravity (discussed more in section 8), but for now we describe states purely in terms of one or the other.

³This is *not* equivalent to the alternative approach of first imposing constraints classically and then quantising. We will see some specific examples in section 7.2.6 where ‘constrain first’ misses some interesting and physically important non-perturbative effects.

Much of the literature which we follow tends to regard the ‘Dirac’ invariant Hilbert space (1.1) as primary, and the unconstrained Hilbert space (and its co-invariant cosets (1.2)) as an auxiliary or kinematic structure used to define the correct Dirac Hilbert space and inner product as described below. While our presentation is technically equivalent, we shift emphasis to put the invariants and co-invariants on an equal footing, regarding them as two special cases among the more general mixtures described in the preceding paragraph.

1.2 Co-invariants, gauge-fixing, and null states

We advocate that for the aims outlined above — a description of states which is local in time in perturbation theory, and which leaves room to account for non-perturbative effects — it is often most useful to describe states as co-invariants. For this, we use a quantum analogue of ‘fixing a gauge’. Specifically, the gauge transformations in question are the redundancies in the description of a state (analogous to the classical freedom that a single spacetime can be defined by many sets of initial data on different Cauchy surfaces); this is distinct from choosing a gauge (e.g., de Donder) in the description of the dynamics or action.

The basic idea is to use the redundancy inherent in co-invariants, namely that a single state can be represented in many ways by choosing a representative of the equivalence class under (1.2). We can ‘gauge-fix’ by allowing only states of a particular chosen form. For example, we might impose the gauge condition $\chi(q) = 0$ by choosing a wavefunction of the form $\psi(q) = \delta(\chi(q))\psi_{\perp}(q_{\perp})$, and the physical content of the state is contained in a wavefunction ψ_{\perp} depending on fewer transverse variables q_{\perp} (chosen so that χ and q_{\perp} together constitute good coordinates for configuration space near $\chi = 0$).⁴ In a semi-classical limit, this can give us a description of states which is (at least approximately) local in time. Dynamics can emerge by changing the choice of gauge-fixing χ (perhaps a different value of some ‘clock’ field), and identifying how the wavefunction ψ_{\perp} changes to keep the state in the same co-invariant coset.

For a good gauge choice, we would like each physical state coset to have a unique representative of the chosen form. This is often difficult or impossible to arrange (a Gribov problem), but it is much easier to find a good gauge in perturbation theory around a given semi-classical background. In this case, the gauge choice eliminates redundancy under infinitesimal diffeomorphisms, but some redundancy under a discrete set of finite non-perturbative (and state-dependent) diffeomorphisms may remain. We propose that such residual diffeomorphisms may be the underlying mechanism for the existence of null states alluded to above (see further comments in section 8.3).

⁴In fact we will often need something slightly more complicated for gravity, such as an additional term proportional to $\delta'(\chi)$ because the Hamiltonian is quadratic in momentum. Additionally, replacing the delta function with a narrow Gaussian is useful for a nice classical limit, since this allows for coherent states which are well-localised in phase space. These points will be addressed in the main text.

1.3 Gravitational inner products

So far, we have only examined the definition of the vector space of physical states. For either \mathcal{H}_{inv} or \mathcal{H}_{co} to be a Hilbert space we need additional structure, namely an inner product. Before discussing specific proposals, we first comment on the general mathematical relationship between invariants and co-invariants and inner products on these spaces.

Because of the duality (1.3) between \mathcal{H}_{inv} and \mathcal{H}_{co} , four different structures become equivalent: an inner product on invariants \mathcal{H}_{inv} , an inner product on co-invariants \mathcal{H}_{co} , a map from co-invariants to invariants $\eta : \mathcal{H}_{\text{co}} \rightarrow \mathcal{H}_{\text{inv}}$, or a map from invariants to co-invariants $\eta^{-1} : \mathcal{H}_{\text{inv}} \rightarrow \mathcal{H}_{\text{co}}$. More precisely, a linear isomorphism of vector spaces η is equivalent to a non-degenerate sesquilinear form on either space; an inner product should also be Hermitian ($\eta^\dagger = \eta$) and positive-definite (or at least semi-definite, $\eta \geq 0$). A consequence of this general structure is that an inner product also allows us to pass between the two ways of defining physical states (1.1) and (1.2). Thus, these are no longer truly different Hilbert spaces (once we have identified an inner product), though one language or the other may be more useful depending on the context.

We can give more physical interpretations to η and its inverse. First, we can think of η as acting on an arbitrary wavefunction in the unconstrained Hilbert space \mathcal{H}_0 (regarded as a representative of a co-invariant coset), mapping it to a solution of the Wheeler-DeWitt equation. So η is roughly like a projection onto gauge-invariant states (though this is not literally true for a non-compact space of gauge transformations like diffeomorphisms). For η to be well-defined on \mathcal{H}_{co} and to map into \mathcal{H}_{inv} , for each constraint \mathcal{H} we should have

$$\eta \mathcal{H} = \mathcal{H} \eta = 0. \quad (1.4)$$

The approach of defining an inner product through such a map goes under the name of ‘refined algebraic quantisation’ [9–11], and η is known as a ‘rigging map’.

The inverse map η^{-1} takes a solution of the Wheeler-DeWitt equation to a co-invariant equivalence class. To make this concrete and practical we would like to represent the outcome as a wavefunction in \mathcal{H}_0 , which means that for each invariant we choose a specific representative of the corresponding coset in \mathcal{H}_{co} . We denote such a concrete representative of η^{-1} by $\kappa : \mathcal{H}_{\text{inv}} \rightarrow \mathcal{H}_0$. This leaves a great deal of freedom in choosing κ , which amounts to a choice of gauge in representing the state. For example, we might be able to choose κ so that it always outputs states satisfying a gauge condition $\chi = 0$ as discussed in section 1.2. We would expect such a map to look something like $\delta(\chi)$ times a Fadeev-Popov determinant factor associated with the gauge-fixing. The details of this construction are fleshed out in the body of the paper. When we use κ to define an inner product on invariants $[[\phi'|\phi]] = \langle \phi' | \kappa | \phi \rangle$, its role is to remove the overcounting from summing over gauge-equivalent wavefunctions. So, we think of κ as a ‘gauge-fixing’ map.

For a given κ to be compatible with a chosen rigging map η , we need it to be a representative of the inverse η^{-1} when thought of as a map into \mathcal{H}_{co} . The condition for this is that κ is a generalised inverse, $\eta \kappa \eta = \eta$:

$$\eta \text{ compatible with } \kappa \iff \eta \kappa \eta = \eta. \quad (1.5)$$

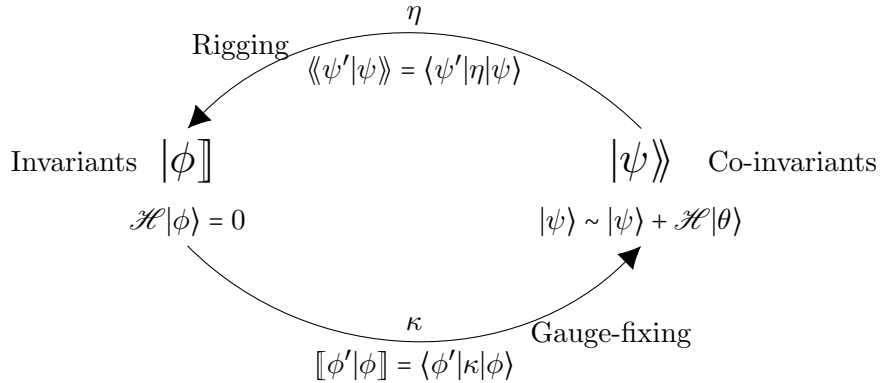


Figure 1. A summary of the relationships between the spaces of invariant and co-invariant states (\mathcal{H}_{inv} and \mathcal{H}_{co}), inner products on these spaces, and maps between them.

One way to read this equation is that $\eta\kappa$ is the identity when acting on invariants, namely on the image of η . Another interpretation is that $\kappa\eta$ acts as a gauge transformation mapping any wavefunction to a different member of the same coset, so $\kappa\eta - 1$ should always yield a pure gauge ‘null state’, which is annihilated by η .

1.4 Group averaging

Having understood the general structure of inner products, we would like a specific candidate for the maps η and/or κ . Perhaps the most well-known proposal is that of DeWitt [7], to use a ‘Klein-Gordon’ inner product on the space of invariants (discussed by Wald [12] for mini-superspace models). This is motivated by the fact that the Hamiltonian constraints \mathcal{H} are quadratic in canonical momenta, so act as second-order differential operators on wavefunctions of metrics analogous to a Klein-Gordon operator; the Klein-Gordon form is then defined in analogy with a similar inner product on the single-particle states of free QFT, defined on solutions to the Klein-Gordon equation. But this has a fatal drawback: it is not positive-definite! The typical remedy is to restrict to a subspace of ‘positive frequency’ states on which the Klein-Gordon form is positive definite, but in general there is no canonical choice for this subspace. In QFT this corresponds to the ambiguity in the choice of a Gaussian vacuum state and the corresponding notion of particle, but for gravity this is not a useful interpretation. Additionally, this restriction typically eliminates states which correspond to sensible, physical classical solutions in the semi-classical limit (for example, cosmologies which expand and re-collapse [13]). With this in mind, we would like a better proposal.

Fortunately, there is a much more satisfactory and well-motivated (but less well-known) suggestion for the physical inner product. This is the ‘group averaging’ procedure [9, 14], which defines the inner product (on the space of co-invariants (1.2)) by averaging over all possible gauge transformations. In the simplest case of a single constraint (which we call H , with 1D gravity in mind) generating the non-compact gauge group \mathbb{R} , we construct a

physical inner product $\langle\langle \cdot | \cdot \rangle\rangle$ as an integral over the gauge group,

$$\langle\langle \psi' | \psi \rangle\rangle = \int_{-\infty}^{\infty} dt \langle \psi' | e^{-iHt} | \psi \rangle = 2\pi \langle \psi' | \delta(H) | \psi \rangle. \quad (1.6)$$

In the language above, $\eta = 2\pi\delta(H)$ is a specific natural choice of rigging map, which generalises the projection onto gauge-invariant states to the case of non-compact gauge groups. It is a particularly nice candidate for an inner product because it is manifestly positive-definite (δ is a positive distribution). Additionally, we will see that it arises naturally both from the path integral and also from the Hamiltonian BRST/BFV formalism, including for constraint algebras with field-dependent structure constants such as gravity. We will also develop a simple understanding of the relationship of group averaging with other candidate inner products (such as Klein-Gordon), and in particular an explanation for their shortcomings (most notably, failure of positivity). For these reasons we advocate for group averaging as the best procedure for defining the inner product on the physical Hilbert space.

1.5 Klein-Gordon and more inner products

Group averaging defines a good physical inner product, and defines an equivalent inner product on a Hilbert space of invariants through the inverse map η^{-1} . However, this inverse is only defined abstractly and is not very practical due to the freedom of choosing a representative wavefunction in a given co-invariant coset. How do we concretely find a representative wavefunction which maps to a given invariant under η ? And how can we compute the inner product between a pair of invariant states directly from the wavefunction? We can answer these questions by finding a concrete κ map, which takes a solution to the Wheeler-DeWitt equation to a specific wavefunction representing the corresponding co-invariant coset.

Unfortunately, we do not expect that there exists any simple κ which inverts η (i.e., satisfying (1.5)). This is essentially due to a Gribov problem: no nice (e.g., local) gauge-fixing condition simultaneously works for all states. Nonetheless, we will be able to find maps κ which invert η on a restricted set of states, in particular to all orders in perturbation theory around a given classical background (or family of backgrounds). For this reason, we will study maps κ which need not satisfy the compatibility condition $\eta\kappa\eta = \eta$ with the group averaging rigging map η : instead, only certain matrix elements of this relation will be satisfied (and even then, perhaps only in perturbation theory). Perhaps the central point of the paper is to explain the circumstances under which this compatibility holds, and the manner in which it is violated.

The Klein-Gordon inner product gives us the prominent example of such a κ map. It arises from a gauge-fixing condition $\chi = 0$ where $\chi(\gamma)$ is a function of spatial metrics γ (independent of their conjugate momenta) at each point of space. As familiar from gauge-fixing procedures in path integrals, it takes the schematic form $\kappa_\chi = \delta(\chi)\Delta_{\text{FP}}$, where Δ_{FP} is a Fadeev-Popov determinant. The $\chi = 0$ condition chooses a hypersurface in superspace (which corresponds in QFT to a Cauchy surface in the target spacetime), and can be written as an integral over that hypersurface. The result is independent of deformations

of the chosen hypersurface, which we can understand as a consequence of gauge invariance (and the job of Δ_{FP}). In the simple case of single Hamiltonian constraint H generating time evolution $\frac{d}{dt} = i[H, \cdot]$, the Klein-Gordon form is given by $\kappa_\chi = \frac{d\chi}{dt} \delta(\chi)$, up to issues of operator ordering (addressed in the main text), so the Fadeev-Popov factor is $\Delta_{\text{FP}} = \frac{d\chi}{dt}$. For gravity, the Hamiltonian is quadratic in momentum, so Δ_{FP} is linear in momentum. This means that the Klein-Gordon form involves first order in derivatives with respect to superspace coordinates γ . Full details will be given in the body of the paper (see also [15]).

In this form, it is clear that we can generalise the Klein-Gordon form by making a different choice of χ which depends not just on superspace variables, but also on their conjugate momenta, and/or on matter fields. An example is the ‘maximal volume gauge’, where χ is the trace of the extrinsic curvature [15]. Another generalisation uses ‘Gaussian averaged’ gauges, where the sharp $\delta(\chi)$ is replaced by a smooth (but narrow) Gaussian. This is a practical choice for semi-classical calculations, since it allows for states which are well-localised in phase space (both χ and its canonically conjugate momentum), corresponding to classical initial data. Some details of these generalisations will be explored in the main text.

Having seen how different maps κ_χ arises from a gauge-fixing choice $\chi = 0$, we can understand the failure of positivity of the Klein-Gordon form (or one of its generalisations). This arises from a bad choice of measure, coming from a possible sign in the Fadeev-Popov determinant $\Delta_{\text{FP}} = \frac{d\chi}{dt}$. From the above, it is clear that positivity of κ depends on whether a given trajectory crosses the gauge-fixing surface $\chi = 0$ in an increasing or decreasing direction: positive-frequency solutions (with positive norm) correspond to increasing χ , and negative-frequency (negative norm) solutions decreasing χ . This is the usual issue that a good measure corresponds to the absolute value of the Fadeev-Popov determinant. An additional potential problem is that a given spacetime may have no Cauchy surface satisfying the gauge condition, or several, and it is hard to find a local gauge-fixing function χ which does not have this drawback. An example is a bang/crunch spacetime which crosses the surface $\chi = 0$ once when expanding and once (in the opposite direction) when re-collapsing; this gives zero Klein-Gordon norm.

This issue is also responsible for κ_χ being incompatible with η . The upshot is that $\eta\kappa\eta = \eta$ is satisfied only for matrix elements where all contributing spacetimes cross the slice $\chi = 0$ once, where we count the number of crossings with signs depending on whether χ is increasing or decreasing. In other words, χ constitutes a ‘good gauge’ for spacetimes which have $\chi < 0$ in the distant past and $\chi > 0$ in the future. We can arrange for this in perturbation theory by a judicious choice of χ , so that $\chi = 0$ selects a unique Cauchy surface Σ in the class of classical background spacetimes of interest, and also χ is increasing on Σ .

Once this issue is taken into account, the compatibility of κ and η is intuitively clearest from a path integral perspective described in the next subsection. From a more algebraic and canonical point of view, the compatibility can be understood by seeing how both η and κ maps emerge from the Hamiltonian BRST/BFV formalism (which is also useful, if not essential, for correctly identifying the appropriate measures).

We have emphasised a perspective where κ is thought of as a map from invariants to

co-invariants, not only as a sesquilinear form on invariants. The reason is that a κ map $\mathcal{H}_{\text{inv}} \rightarrow \mathcal{H}_{\text{co}}$ is itself useful, since it gives us a way to construct gauge-fixed wavefunctions as described in section 1.2. The idea is that co-invariant states defined by ‘initial data’ on a Cauchy surface selected by the gauge condition $\chi = 0$ are constructed as the image of the map κ_χ . Note that we can construct states in this way even if χ is not a globally good gauge condition (especially in perturbation theory), treating κ as a tool for generating an ansatz while relying on the group averaging η to provide the correct physical inner product.

1.6 Gravitational path integrals

Many recent successes in quantum gravity have been cast in the formalism of the gravitational path integral. This approach is not at odds with the canonical formalism we are using here: on the contrary, they are complementary approaches which are useful for emphasising different aspects of the physics. In particular, a good understanding of the ideas of canonical quantum gravity is invaluable for providing well-justified Hilbert space interpretations of path integral calculations. Conversely, the path integral provides useful intuition for the sort of canonical constructions we have been describing, as well as practical methods for semi-classical calculations.

States and group averaging

Consider first a path integral over spacetimes bounded by a pair of Cauchy surfaces Σ_1 and Σ_2 , with fixed field configurations on those surfaces: that is, fixed spatial metrics γ_1, γ_2 along with matter fields. In a non-gravitational theory, such a path integral (with fixed spacetime geometry) computes the matrix elements of a time-evolution operator, relating wavefunctionals depending on fields on different Cauchy surfaces. Most simply, if spacetime is a product of space Σ and an interval of time, the path integral computes matrix elements of e^{-iHt} , where H is the Hamiltonian of the theory on Σ . But for gravity, the metric is a dynamical field so we also integrate over all possible spacetime metrics g between $\Sigma_{1,2}$ (with induced metrics $\gamma_{1,2}$ on the boundaries $\Sigma_{1,2}$), modulo diffeomorphisms which leave $\Sigma_{1,2}$ invariant. Since time-evolution (as well as changes of coordinates on a fixed spatial slice) is part of the gauge symmetry of diffeomorphisms, the integral over such geometries is precisely an integral over the gauge group.⁵ So, the path integral is computing the matrix elements $\langle \gamma_2 | \eta | \gamma_1 \rangle$ of the group-average rigging map introduced in section 1.4.⁶ For a one-dimensional (or mini-superspace) theory, the integral over metrics becomes an integral over the proper time t ; since the integral over all other fields gives the time-evolution operator e^{-iHt} , we recover precisely (1.6). We also learn that a fixed spatial metric γ_1 (or more general finite time boundary conditions) should be interpreted as states $|\gamma_1\rangle$ with delta-function wavefunctions $\langle \gamma | \gamma_1 \rangle = \delta(\gamma - \gamma_1)$ in the *unconstrained*

⁵More precisely, this integrates over diffeomorphisms of spacetime, modulo the diffs which leave the spatial slice Σ_1 invariant.

⁶In a situation with fixed spatial boundary conditions (such as asymptotically AdS), we are referring to an integral where Σ_1 and Σ_2 are required to coincide at the boundary. If we instead require the initial and final Cauchy surfaces go to the boundary at different times, we are not simply computing an inner product, but also inserting a physical time-evolution operator.

Hilbert space, which we can regard as representatives of a co-invariant equivalence class $|\gamma_1\rangle\rangle$. This path integral computes the physical (group-averaging) inner product $\langle\langle\gamma_2|\gamma_1\rangle\rangle$ on co-invariant states.

Note that this interpretation as a gauge-invariant inner product on co-invariants only makes sense if the path integral constructs solutions to the constraints [16]. Concretely, $\Psi(\gamma) = \langle\gamma|\eta|\gamma_1\rangle$ should solve the Wheeler-DeWitt equation, or more abstractly $|\eta\gamma_1\rangle = \eta|\gamma_1\rangle\rangle$ should be in the Hilbert space of invariants. This requires us to integrate over spacetimes where Σ_1 can lie to the future of Σ_2 or to its past (or a mixture of the two): in ADM language, we integrate over both positive and negative lapse. Integrating only over positive lapse would result in a ‘propagator’ for the constraints (solving $H\Psi(\gamma) = \delta(\gamma - \gamma_1)$ instead of $H\Psi(\gamma) = 0$). See further comments in [17].

This illustrates that any boundary condition on a fixed finite-time Cauchy surface can be interpreted as a co-invariant state, and also shows how path integrals can define invariant wavefunctions as a function of boundary conditions. The latter is not limited only to the group-averaging we just saw, and many familiar classes of states are prepared by different path integrals. An important example are states defined by asymptotic boundary conditions in the distant past, essentially the same as scattering ‘in’ states (or similarly ‘out’ states defined in the future). These can be prepared by path integrals over only forward time evolution (positive lapse), except in a limit for which only very large times contribute, so the WDW equation is still satisfied (the would-be source which appeared above is pushed off to infinity). An example would be out states defined at future infinity in de Sitter space (discussed for JT gravity in [5]). Such constructions could also be generalised to allow some Euclidean evolution, for example to prepare a vacuum state in asymptotically AdS spacetimes, or the Hartle-Hawking state. The wavefunction $\Psi(\gamma)$ of such a state is interpreted in our language as the canonical pairing $\langle\langle\gamma|\psi\rangle\rangle$ between co-invariants and invariants.

Cutting path integrals, gauge-fixing, and κ maps

Let’s now consider how we recover a Hilbert space interpretation from a path integral. In ordinary QFT, we do this by cutting spacetime into two along a Cauchy surface Σ , producing a Hilbert space \mathcal{H}_Σ which can be informally thought of as wavefunctions $\psi(q)$ where q denotes the field configuration on Σ . The integral over one half of spacetime (with fixed field configuration q on Σ) defines a wavefunction $\psi_1(q)$ of a ‘ket’ state $|\psi_1\rangle$, the integral over the other half similarly defines a ‘bra’ state $\langle\psi_2|$, and integrating over fields q on Σ then sews the two halves together. This last integral is interpreted in Hilbert space language as a completeness relation, the inverse to the inner product on Σ . Arguably, such a factorisation of path integrals is the main purpose of a Hilbert space: it embodies locality (since the states are functions of local data on Σ) and positivity (since the inner product is positive-definite). This allows us to define local operators on Σ and assign sensible probabilistic interpretations to their expectation values. For gravity it is not so simple: diffeomorphism invariance means we cannot simply cut along a ‘ $t = 0$ ’ Cauchy surface Σ , because gauge transformations move the physical location of Σ .

To illustrate the difficulty, let's consider cutting the above path integral between surfaces $\Sigma_{1,2}$ computing the group-average $\langle\langle \gamma_2 | \gamma_1 \rangle\rangle = \langle \gamma_2 | \eta | \gamma_1 \rangle$. Suppose we naïvely follow the usual procedure of integrating over all field configurations (spatial metrics γ) on an intermediate surface Σ , as well as all metrics g_1 between Σ_1 and Σ (with fixed boundary metrics γ_1, γ , and modding out by diffeomorphisms which are trivial on these boundaries) and similarly g_2 between Σ and Σ_2 . The result is not the same as the integral over all metrics which gave us matrix elements of η : in fact, the result additionally integrates over all possible embeddings of the Cauchy surface Σ in a given spacetime geometry (including all possible choices of coordinates on Σ itself). We have overcounted by the volume of a residual gauge group: diffeomorphisms modulo those which are trivial on Σ (see (2.39) of [15] for a similar discussion). This construction computes matrix elements of η^2 , which does not make sense for non-compact (and field-dependent) groups of gauge transformations, even for $\eta = 2\pi\delta(H)$ in the simplest 1D example given above.

The remedy is obvious from this perspective: we must implement a gauge-fixing which identifies Σ uniquely in some diffeomorphism-invariant way. As we described in section 1.5 above, this is precisely what the maps κ (such as the Klein-Gordon inner product) should do. The equation $\eta\kappa\eta = \eta$ given in (1.5) is a manifestation of this process of cutting and gluing path integrals with a good gauge choice, shown in figure 2. Each η on the left represents the path integral between Σ_1 and Σ or between Σ and Σ_2 , while κ is an insertion of a gauge-fixing $\delta(\chi)$ which chooses a physical embedding of Σ uniquely, along with an appropriate (positive) measure factor Δ_{FP} . Then the product $\eta\kappa\eta$ (sewing along Σ) recovers η on the right hand side, the integral over spacetimes between Σ_1 and Σ_2 . Thus, we can think of κ as giving a ‘completeness relation’ dual to the group average inner product η , taking gauge redundancy into account. But as emphasised already, this relation can be violated by spacetimes where κ does not select a unique Σ .

1.7 Inner products in the BRST formalism

The BRST formalism provides a very general framework for systems with constraints. It naturally incorporates both the invariant and co-invariant states described above (and arbitrary mixtures of the two). It also provides a mechanism to describe the inner products introduced above (including the particular cases of group averaging and the Klein-Gordon form) along with a physical interpretation in terms of gauge-fixing. In doing so, it provides a certain rigidity absent from our considerations so far, removing much of the ambiguity in choosing an inner product, and providing further justification for the group-averaging prescription. Finally, it is often a powerful technical tool for identifying the correct gauge-invariant form of inner products and observables.

In this formalism, for each constraint \mathcal{H} we introduce corresponding fermionic ghost variables c, b , which act as raising and lowering operators respectively on a two-level Hilbert space spanned by $|\uparrow\rangle, |\downarrow\rangle$. The central object is a BRST operator $Q = \int c\mathcal{H} + \dots$ built from the ghosts and constraints, with \dots determined by the requirements of nilpotency $Q^2 = 0$ and Hermiticity $Q^\dagger = Q$. The physical Hilbert space is then given by the cohomology of Q , $\ker Q / \text{im } Q$ (at ghost number zero). This gives different ways to represent physical states depending on the state of the ghosts. At the ‘lower level’ $|\downarrow\rangle$ the constraints are important

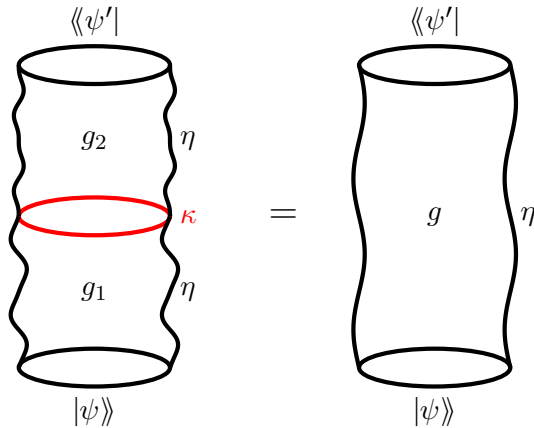


Figure 2. The generalised inverse relation $\eta\kappa\eta = \eta$ between group averaging and gauge-fixing maps has a geometric interpretation in terms of cutting and gluing spacetimes in the path integral. On the left, we integrate over all geometries g_1 and g_2 , glued together with an operator κ which implements a gauge condition fixing the red cutting surface, giving matrix elements of $\eta\kappa\eta$. If this is a good gauge (selecting a unique intermediate Cauchy surface in every geometry), then the result is equivalent to the integral over all geometries g which computes η .

for determining $\ker Q$, giving rise to the invariant Hilbert space. At the ‘upper level’ $|\uparrow\rangle$ the constraints instead determine the denominator $\text{im } Q$, giving us cosets corresponding to co-invariant states. So, the myriad of possible ways of imposing constraints correspond to choices of ghost state in a BRST cohomology class.

When we calculate inner products in this formalism, we immediately recover the natural pairing $\llbracket \psi | \psi' \rrbracket$ between invariants and co-invariants. However, inner products between two states of the same type are more subtle: a straightforward approach gives us an indeterminate form $\infty \times 0$, where ∞ comes from the original phase space variables and 0 from the ghosts. We can interpret this as arising from a failure to gauge-fix, with ∞ coming from the volume of the unfixed gauge group. Following [18–20], the key idea to fix this is that the degeneracy can be lifted by inserting a trivial operator $\exp[Q, \Psi]_+$ (the exponential of a BRST-exact operator), physically equivalent to the identity. Here Ψ is a fermionic operator of ghost-number -1 which we are free to choose, called the ‘gauge-fixing fermion’. An appropriate choice of Ψ recovers each of the inner products discussed above. The ghost-free part of $[Q, \Psi]_+$ gives the generator of a gauge transformation, while the other terms give us insertions of ghosts b (for co-invariant states) or c (for invariant states) along with gauge-fixing Jacobians. The result is invariant under infinitesimal changes of Ψ (the Fradkin-Vilkovisky theorem [21]).

However, the result is not totally independent of Ψ : there is a loophole to the Fradkin-Vilkovisky theorem, as the result can jump for finite changes when passing through a singular value (most notably $\Psi = 0$, where we get the original indeterminate result). This is the cause (from the BRST perspective) of the signs discussed in section 1.5, including indefiniteness of the Klein-Gordon inner product.

More specifically, we use an extended Hamiltonian form of the BRST formalism known

as the Batalin-Fradkin-Vilkovisky method [21–24]. In this BFV-BRST formalism, we work in an extended phase space which retains the Lagrange multipliers which impose constraints — the lapse N and shift N_{\perp}^i in gravity (or the timelike component of the gauge field A_0 in gauge theories) — along with their conjugate momenta Π . These extra variables are ultimately removed by also including an extra constraint that $\Pi = 0$, but nonetheless prove helpful to include both technically and conceptually. In particular, the freedom in the gauge-fixing fermion Ψ can be used to choose a gauge, with the gauge-fixing condition coming from the equation of motion for Π (so that integrating out Π and other momenta reproduces the Fadeev-Popov path integral). For example, we can choose Ψ to compute group average inner products using various familiar coordinate conditions (e.g., synchronous or de Donder gauges) including Gaussian averaged versions. The further technical virtue of this formalism is that it naturally gives us a group averaging procedure for rather general algebras of constraints, including those with field-dependent structure constants as arise in gravity (discussed further in section 8).

1.8 Perturbative and non-perturbative physics in the semi-classical limit

With the exception of simple toy models, we must face the fact that exact calculations in gravity are essentially impossible. This is not only a matter of technical difficulty: the non-renormalisability of GR means that new physics is inevitably necessary in the UV, which surely replaces spacetime with some more fundamental description.⁷ So, our tools are only useful if we can practically apply them in a semi-classical approximation. In particular, we would like to sketch how the ideas we have discussed can give a semi-classical gravitational Hilbert space satisfying the criteria listed at the start of this introduction.

At leading order in G_N we would like states to be described by local data on a specific Cauchy surface in a classical background spacetime. The Wheeler-DeWitt prescription of invariant states does not achieve this (at least by itself), because such wavefunctions are superpositions over all possible Cauchy surfaces, obscuring locality. However, we do get such a local description from co-invariant wavefunctions with a gauge-fixing ansatz as sketched in section 1.2 (perhaps obtained as the image of some κ map as mentioned at the end of section 1.5). We can think of this as a quantum version of initial data on a Cauchy surface specified by a gauge condition $\chi = 0$. This correspondence to classical initial data is sharpest if we choose coherent wavefunctions (which means we do not impose $\chi = 0$ as a sharp delta function), for example using Gaussian averaged versions of the κ map, so that our states are well-localised in phase space.⁸ Working with such gauge-fixed states around a

⁷Holographic duality achieves this, but we still face the challenge of extracting local bulk physics from the available dual observables, which presumably requires understanding gravitational variables of some sort.

⁸A similar possibility is to specify a state by an invariant wavefunction, but only giving the data of the wavefunction (and perhaps a finite number of its derivatives) evaluated on the gauge-fixed surface $\chi = 0$ in superspace, in the hope that this picks out a unique solution to the Wheeler-DeWitt equation. This is closely related to our proposal, differing by insertions of the measure factor Δ_{FP} (which may depend on data away from $\chi = 0$, in particular requiring derivatives of the invariant wavefunction normal to χ). So, this alternative should be satisfactory in perturbation theory (and locally in superspace), but the co-invariant perspective is more robust, and useful in particular for understanding non-perturbative effects (in the group

classical background achieves the first two aims in our wish list: a description which reduces to local QFT when $G_N \rightarrow 0$, with a systematic perturbation theory to compute corrections order-by-order in G_N . In particular, group averaging supplies the inner product, and the BRST formalism provides the tools for computing this perturbatively in any desired gauge.

For a pair of states which are perturbations around the same background and defined on the same gauge-fixed Cauchy surface, the main semi-classical contribution to the inner product comes from short time evolution (expanding around zero lapse and shift). It is this piece that we expect to give back the local QFT inner product (and perturbative corrections will still be local, albeit containing higher order spatial derivatives). But there can also be semi-classical saddle-points which involve a finite time evolution, and these are responsible for several virtues of this approach.

Most simply, if we have a pair of states defined for two different gauge slices on the same classical background, there will be a real Lorentzian saddle point contributing to the inner product (corresponding to the piece of spacetime between the respective Cauchy surfaces). This inner product essentially becomes the matrix elements of a time-evolution operator: the evolution is purely a gauge transformation in gravity, but in this way we can nonetheless recover a notion of (approximately) local dynamics.

Additionally, the group average path integral can have saddle-points with complex metrics. The defining contour runs over real Lorentzian metrics, but this may be deformed to a steepest-descent contour which passes through a saddle-point with complex-valued metric. Such a saddle must be exponentially suppressed (its action will have positive imaginary part), but can nonetheless describe interesting physical effects. In particular, they can give the leading contribution to processes (such as decay by quantum tunneling) which are classically completely forbidden. If we allow our integral to include certain mild relaxations of Lorentzian geometries, these processes could potentially include topology change (as discussed in section 8). More generally, non-trivial saddle-points provide contributions to the inner product which are qualitatively different from the perturbative piece described above because they are completely non-local in space. We speculate that such corrections could help to explain the null state phenomenon mentioned in the introduction.

1.9 Paper outline

The remainder of this paper explains the details of the ideas presented above in the simplest case of one-dimensional (or minisuperspace) theories of gravity with a single constraint, which we introduce in section 2. In section 3 we describe different approaches for imposing constraints to give a space of physical states, and examples of interesting states in each approach. We then turn to the inner product in section 4, including explicit calculations in a particularly simple class of examples with one-dimensional target space. A brief discussion of physical operators (in both Dirac and BRST formalisms) is given in section 5, which is mostly a preliminary for understanding the BFV-BRST formalism in section 6. We then discuss applications in the semi-classical limit in section 7. We conclude with discussion in section 8.

average inner product, for example).

Appendices [E](#) and [F](#) connect the ideas in this paper to analogous concepts in the string worldsheet and abelian gauge theory respectively. Other appendices contain some additional technical details.

2 One-dimensional (or mini-superspace) gravity

For simplicity and clarity we will spend most of our time illustrating the ideas in one-dimensional gravity (or mini-superspace) models. This is sufficient to discuss the gauging of reparameterisations of time, or the Hamiltonian constraint. While a full higher dimensional theory of gravity adds significant technical complication, the conceptual issues associated with defining Hilbert spaces and inner products are well illustrated by these simpler one-dimensional models.

2.1 Coupling quantum mechanics to one-dimensional gravity

To motivate the following, consider a ‘matter’ theory in one dimension with Lagrangian $L(q, \dot{q})$ (where q denotes a collection of n configuration space variables), and minimally couple it to one-dimensional gravity. The only component of a one-dimensional metric $g = -N(t)^2 dt^2$ is the ‘lapse’ $N(t)$, or equivalently the derivative $N = \frac{d\tau}{dt}$ of proper time τ with respect to coordinate time t . (Another possible name and notation for N is an ‘einbein’ e , defined so that $e(t)dt$ is a unit one-form.) We construct a diffeomorphism-invariant Lagrangian \mathbf{L} by ‘minimal coupling’ of the matter to the dynamical metric,

$$\mathbf{L} = NL(q, N^{-1}\dot{q}). \tag{2.1}$$

There is no pure gravitational ‘Einstein-Hilbert’ term since there is no curvature in one dimension; the only possible term involving only N is a ‘cosmological constant’ ΛN , but this can be absorbed into a constant shift of L . The usual Legendre transform gives us the Hamiltonian

$$\mathbf{H} = NH(q, p), \tag{2.2}$$

where $H(q, p)$ is the ‘matter’ Hamiltonian constructed from $L(q, \dot{q})$ in the usual way. Since the Lagrangian \mathbf{L} does not depend on derivatives of N , we have the primary constraint that its conjugate momentum vanishes, $\Pi = p_N = 0$, and the Poisson bracket of \mathbf{H} with Π generates the secondary constraint $H = 0$.

The upshot is that the gravitational theory is constructed from the ‘matter’ theory by imposing the Hamiltonian constraint $H = 0$. Classically, this means that states are given by the ‘symplectic quotient’ of phase space, meaning the space of zero energy trajectories (points on the constraint surface $H(q, p) = 0$, with any two points along the same flow generated by H identified). Quantum mechanically, we can take essentially any ordinary ‘matter’ quantum mechanics theory with Hilbert space \mathcal{H}_0 , inner product $\langle \cdot | \cdot \rangle$, and Hermitian Hamiltonian H (the only assumption we need to make is that H has continuous spectrum in a neighborhood of zero energy) and couple it to 1D gravity by imposing $H = 0$ at the quantum level. The main point of this paper is to review and explore the details of how we do this.

For much of the paper we do not include the lapse as a separate degree of freedom, regarding N as simply a Lagrange multiplier imposing the $H = 0$ constraint. Alternatively, we can interpret this as eliminating the $\Pi = 0$ primary constraint by fixing the gauge $N = 1$. However, for some purposes it is useful to use a ‘non-minimal’ formulation which explicitly retains N , Π and the constraint $\Pi = 0$; we return to this in section 6.

The constraint H generates a gauge symmetry, but note that in the Hamiltonian formalism this is not quite the full group of diffeomorphisms of time (or spacetime in higher dimensions) that we might expect from the Lagrangian formulation. Rather, it generates only the part that acts on the $t = 0$ slice where the wavefunction is defined. More precisely, it gives the group of diffeomorphisms modulo the diffs that act as the identity at $t = 0$. In one dimension, this becomes simply the one-dimensional group of rigid time translations generated by H .

Finally, note that we ‘quantise first’, constructing an unconstrained quantum theory and then imposing $H = 0$ quantum mechanically, rather than ‘constraining first’, which means quantising the constrained symplectic quotient phase space. These two operations do not commute in general: in particular, ‘constrain first’ can eliminate interesting and important non-perturbative effects (we discuss an example in section 7.2.6).

2.2 Examples: sigma models

For a very concrete set of examples we can consider a sigma-model:

$$\mathbf{L} = \frac{1}{2N} G_{ab}(q) \dot{q}^a \dot{q}^b - NV(q). \quad (2.3)$$

Here G_{ab} is the metric on some ‘target space’ manifold \mathcal{M} , which is non-degenerate but can have any signature; in particular, it is not required that H is bounded from below to get a physically sensible theory. V is a ‘potential’ on this manifold. This is a general bosonic model with a Lagrangian quadratic in velocities and time-reversal symmetry (relaxing this constraint allows an additional a ‘background magnetic field’ term $A_a(q)\dot{q}^a$). Classically, the Hamiltonian is

$$H = \frac{1}{2} G^{ab}(q) p_a p_b + V(q). \quad (2.4)$$

The quantum Hamiltonian has an operator ordering ambiguity; for definiteness we make the choice $H = -\frac{1}{2}\nabla^2 + V$ where ∇^2 is the Laplacian on target space \mathcal{M} with metric G . Then, the matter Hilbert space is the square integrable (scalar) functions $L^2(\mathcal{M})$ of \mathcal{M} with the measure $d^n q \sqrt{|\det G|}$.⁹ It is natural to also consider a quantum correction ξR to the potential, proportional to the curvature of target space; in particular, the ‘conformal coupling’ value $\xi = -\frac{1}{8} \frac{n-2}{n-1}$ may be preferred [25, 26]. In any case, we can always absorb this in the definition of V .

⁹A subtlety with this choice is that p_a is not equal to $-i\frac{\partial}{\partial q^a}$, since this is not Hermitian. Instead we should take $p_a = -i\frac{\partial}{\partial q^a} - \frac{i}{4}(\partial_a \log G)$ (with $G = |\det G|$). More generally, for any vector field v on target space, the operator $-iv^a \partial_a - \frac{i}{2}(\nabla_a v^a)$ (obtained by acting with v^a and covariant derivative ∇_a and averaging over the two orders) is Hermitian; we get the above by taking $v = \partial_a$. An alternative takes wavefunctions to be ‘half-densities’ so that $|\psi|^2$ transforms as a volume form under changes of target space coordinates. This gives $p_a = -i\frac{\partial}{\partial q^a}$. In either case, the operator ordering for the kinetic term is $G^{-\frac{1}{4}} p_a G^{ab} G^{\frac{1}{2}} p_b G^{-\frac{1}{4}}$.

This set of models give reasonable analogues to higher dimensional theories of gravity since the Hamiltonian constraint is quadratic in momenta. For this, we interpret the target space \mathcal{M} as Wheeler’s ‘superspace’, which is the collection of all possible spatial geometries (Riemannian manifolds modulo diffeomorphisms). In fact, we obtain a theory of this kind by a mini-superspace truncation of a higher-dimensional theory. This means restricting the integral over metrics to a finite-dimensional family of homogeneous spatial metrics $\gamma_{ij}(q)dx^i dx^j$ preserving some symmetries (parameterised by the coordinates q). Making a metric ansatz $ds^2 = -N(t)^2 dt^2 + \gamma_{ij}(q(t))dx^i dx^j$ (along with other fields respecting the same symmetry, depending only on time), we arrive at a one-dimensional theory of gravity precisely of the form (2.4). Typically the target space metric G has Lorentzian signature, with the timelike direction corresponding roughly to the volume of space. An example is an FLRW cosmology, for which the Hamiltonian constraint $H = 0$ is the Friedmann equation and the scale factor a is a timelike coordinate in superspace.

2.3 One dimensional gravity vs worldline QFT

In cases where the target space metric G is Euclidean or Lorentzian, the model (2.4) is also closely related to free bosonic QFT on \mathcal{M} where the potential becomes a spacetime dependent mass squared $V = \frac{1}{2}m^2$. We then think of our quantum mechanics as living on the worldline of particles, and interpret its wavefunction as a field on target space. The QFT or target space action is given by the expectation value of H ,

$$S_{\text{target}}[\Phi] = \int d^n q \sqrt{|\det G|} \left[\frac{1}{2} G^{ab}(q) \partial_a \Phi^*(q) \partial_b \Phi(q) + V(q) |\Phi(q)|^2 \right]. \quad (2.5)$$

The Hamiltonian H of our one-dimensional theory becomes the equation of motion $H\Phi = 0$ for the quantum field Φ . We have written a theory of a complex scalar; for a real scalar we can restrict to states invariant under time-reversal of the worldline theory (which is unrelated to target-space time reversal). We will discuss this worldline or ‘first quantised’ formulation of QFT later, in particular emphasising that the perspective of a one-dimensional theory of gravity naturally leads to a different Hilbert space than QFT (see also [27]).

3 States: invariants and co-invariants

3.1 States and the Hamiltonian constraint

Starting from the unconstrained Hilbert Space \mathcal{H}_0 , there are in fact two natural ways to construct a physical gauge-fixed Hilbert space by imposing the constraint $H = 0$.

The first is the ‘Dirac’ approach [6], which is perhaps more familiar and standard in the literature. For this we restrict to ‘gauge-invariant states’ which are left unchanged by time translation e^{-iHt} . These are annihilated by the constraint H , so they are solutions to the Wheeler-DeWitt equation $H|\psi\rangle = 0$. We denote these *invariants* by $|\psi\rangle$, giving a space of states

$$V_{\text{inv}} = \ker(H) = \{|\psi\rangle : H|\psi\rangle = 0\}. \quad (3.1)$$

Since the spectrum of H is continuous at zero, strictly speaking these invariants are not elements of \mathcal{H}_0 since they are not normalisable (we comment on the precise mathematical

definition of V_{inv} in the next subsection). Energy eigenstates $|\psi_E\rangle$ are δ -function normalised, $\langle\psi_{E'}|\psi_E\rangle \propto \delta(E - E')$, so the \mathcal{H}_0 inner product of two invariants contains an infinite factor $\delta(E = 0)$. The non-normalisability means that we do not get a sensible inner product on V_{inv} simply by restriction of the \mathcal{H}_0 inner product to the invariants, so V_{inv} is (at this point) only a vector space of states, not a Hilbert space. We will discuss several ways to put an inner product on V_{inv} later.

The second natural construction of gauge-invariant states is the space of *co-invariants*, where the name (following [8]) reflects that these are dual to the invariants (in the sense of linear algebra). For this, instead of imposing a restriction on the states we impose an equivalence relation under action of the constraints:

$$V_{\text{co}} = \mathcal{H}_0 / \text{Im}(H), \quad (3.2)$$

so states in V_{co} (which we denote by $|\psi\rangle\rangle$) are cosets

$$|\psi\rangle\rangle = \{|\psi\rangle + H|\chi\rangle : |\chi\rangle \in \mathcal{H}_0\}. \quad (3.3)$$

In other words, $|\psi\rangle\rangle$ is an equivalence class of states under gauge transformations generated by H . We have written this in terms of infinitesimal gauge transformations, though we get the same by imposing equivalence under finite transformations $|\psi\rangle \sim e^{-iHt}|\psi\rangle$ for any t . Like the invariants, V_{co} is for now only a vector space of states, and we will later discuss methods for equipping it with an inner product.

To see that the space of invariant states V_{inv} is dual to the co-invariants V_{co} , note that there is a natural sesquilinear pairing $[[\phi|\psi\rangle\rangle$ between an invariant $|\phi\rangle$ and a coset $|\psi\rangle\rangle$ (or equivalently the complex conjugate $\langle\langle\psi|\phi\rangle\rangle$). This pairing is simply the inner product in \mathcal{H}_0 , $[[\phi|\psi\rangle\rangle := \langle\phi|\psi\rangle$ where we may take $|\psi\rangle$ to be any representative of the coset; this is well-defined since H is Hermitian on \mathcal{H}_0 and $|\psi\rangle$ is invariant. This gives an anti-linear map $|\phi\rangle \mapsto [[\phi|\cdot\rangle\rangle$ from V_{inv} to V_{co}^* , the space of linear functionals $V_{\text{co}} \rightarrow \mathbb{C}$, which is bijective (every functional is given by pairing with some invariant, and every non-zero invariant gives a non-zero functional). Hence we have an anti-linear isomorphism $V_{\text{inv}} \simeq V_{\text{co}}^*$ between invariants and the dual of co-invariants. In fact, in a more mathematically precise treatment this isomorphism can be taken to be the *definition* of V_{inv} , as we sketch now.

3.2 Technical comments

Since there are no normalisable states annihilated by H (under the assumption of continuous spectrum at zero energy), a precise definition of V_{inv} requires an enlarged space of distributional states. For this, we first choose a dense subspace $\Phi \subset \mathcal{H}_0$ of ‘test functions’ (equipped with an appropriate topology). For example, in a sigma model as in section 2.2 we might take Φ to be the smooth compactly supported functions on target space \mathcal{M} (though this is more restrictive than necessary). Then distributional states are defined as elements in the dual space Φ^* of continuous linear functionals on Φ , though in fact it will be more convenient to use the complex-conjugate space $\bar{\Phi}^*$ of antilinear functionals. So, $|\phi\rangle \in \bar{\Phi}^*$ is defined via the map $|\psi\rangle \mapsto \langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$ taking $|\psi\rangle \in \Phi$ to \mathbb{C} .

If H maps test functions to test functions (and is continuous with respect to the chosen topology on Φ), then using its self-adjointness we can define H on $\bar{\Phi}^*$ via duality: $H|\phi\rangle$

is defined by the antilinear map taking $|\psi\rangle \in \Phi$ to $\langle\phi|H|\psi\rangle^*$. Invariants are then defined as distributions annihilated by H , $V_{\text{inv}} = \ker H \subset \bar{\Phi}^*$; these are the antilinear functionals which vanish on the image of H . To get a nice pairing between invariants and co-invariants, we must construct V_{co} not from all of \mathcal{H}_0 but only from test functions, so $V_{\text{co}} = \Phi/(\text{im } H)$. Thus, in the equation $\llbracket\phi|\psi\rangle\rangle := \langle\phi|\psi\rangle$ we should read the right hand side as a pairing between $|\phi\rangle \in V_{\text{inv}} \subset \bar{\Phi}^*$ and $|\psi\rangle \in \Phi$. From these definitions, an alternative characterisation of V_{inv} is the complex conjugate of the dual of co-invariants \bar{V}_{co}^* (continuous antilinear functionals on V_{co} , with continuity defined by the quotient topology on V_{co}). In other words, we can use the isomorphism at the end of the last subsection to *define* V_{inv} .

There is a lot of freedom in choosing Φ and the choice certainly has an impact on various of the constructions below. But once we have chosen an inner product and passed to the completion to get the physical Hilbert space, many of the precise details will not matter. So, we will not distract ourselves too much with these details, only making a few comments in footnotes.

See [28] for a much more detailed discussion of technical issues in systems with a single constraint.

3.3 Invariants and co-invariants in BRST quantisation

These two spaces of states arise naturally in the framework of BRST quantisation. Associated with the constraint $H = 0$ there is a single pair of (Hermitian) ghosts b, c with anti-commutator $[b, c]_+ = 1$, represented on a two-dimensional ghost Hilbert space \mathcal{H}_g :

$$\mathcal{H}_g \text{ spanned by } |\downarrow\rangle, |\uparrow\rangle, \text{ with } c|\downarrow\rangle = |\uparrow\rangle, b|\uparrow\rangle = |\downarrow\rangle, b|\downarrow\rangle = c|\uparrow\rangle = 0. \quad (3.4)$$

Physical states are then described by the cohomology of the BRST operator $Q = cH$ acting on $\mathcal{H}_0 \otimes \mathcal{H}_g$:

$$\mathcal{H}_{\text{BRST}} = \frac{\ker Q}{\text{im } Q}, \quad Q = cH. \quad (3.5)$$

This Hilbert space further decomposes into sectors of fixed ‘ghost number’ $N_g = cb - \frac{1}{2}$ (with the constant chosen so that $N_g^\dagger = -N_g$), so states $|\psi\rangle \otimes |\downarrow\rangle$ and $|\psi\rangle \otimes |\uparrow\rangle$ have ghost number $-\frac{1}{2}$ and $+\frac{1}{2}$ respectively for $|\psi\rangle \in \mathcal{H}_0$. The BRST operator has ghost number one, with $Q|\psi\rangle \otimes |\downarrow\rangle = H|\psi\rangle \otimes |\uparrow\rangle$ and $Q|\psi\rangle \otimes |\uparrow\rangle = 0$. At the ‘lower level’ of cohomology the image of Q is trivial (only 0) so physical states are simply given by the kernel of Q , which is the space of invariants V_{inv} . At the ‘upper level’ Q annihilates everything, so the cohomology consists of cosets with respect to the image of Q , which is the space of co-invariants V_{co} :

$$\mathcal{H}_{\text{BRST}} \simeq V_{\text{inv}} \otimes |\downarrow\rangle \oplus V_{\text{co}} \otimes |\uparrow\rangle. \quad (3.6)$$

A physical Hilbert space should consist only of states at some fixed ghost number, so we should take either invariants or co-invariants, not both.

Note that for more complicated theories with multiple constraints, we have many choices in how to implement them: all invariant (à la Dirac), all co-invariant, or a mixture (the typical approach in gauge theories). Correspondingly, in BRST we have a single ghost for every constraint. We comment on these choices for gravity in the discussion 8, along with

the additional technicalities (non-abelian constraint algebras and field dependent structure constants) that arise in BRST with several constraints. We will focus throughout on ideas which readily generalise to this more interesting context.

This BRST Hilbert space has a natural inner product induced from inner products on \mathcal{H}_0 and on \mathcal{H}_g . The latter is determined (up to normalisation) by Hermiticity of the ghost operators b, c to be $\langle \uparrow | \downarrow \rangle = \langle \downarrow | \uparrow \rangle = 1$, with $\langle \uparrow | \uparrow \rangle = \langle \downarrow | \downarrow \rangle = 0$. This means that states in the lower level $|\downarrow\rangle$ pair with states in the upper level $|\uparrow\rangle$ with the opposite ghost number. The result is simply constructed from the above sesquilinear pairing $[\langle \cdot | \cdot \rangle], \langle \langle \cdot | \cdot \rangle \rangle$ between invariants and co-invariants. It does not induce an inner product at any fixed ghost number and it is not positive-definite, so is not a candidate for the physical inner product of the constrained theory. Any pairing between co-invariants should naturally be accompanied by a b ghost insertion, and a pairing between invariants should come with a c ghost; we will see later how these arise from the residual diffeomorphism symmetry.

3.3.1 Non-minimal BRST

We will later make use of the non-minimal formalism [24] where we retain the lapse N as a dynamical variable, along with the additional constraint that its canonical momentum vanishes, $\Pi = 0$. Before imposing constraints, states are wavefunctions $\psi(q, N)$ depending on N (which we allow to take any real value, allowing for proper time to run in the opposite direction to coordinate time).¹⁰ In the BRST formalism, an additional pair of ghosts \hat{b}, \hat{c} are associated with the extra constraint. The ghost number is $N_g = cb - \hat{b}\hat{c}$, and the BRST charge becomes $Q = cH + \hat{c}\Pi$. The ghost Hilbert space is then four-dimensional, with a state $|\downarrow\downarrow\rangle$ at ghost number $N_g = -1$ annihilated by both b, \hat{b} , two states $|\uparrow\downarrow\rangle = c|\downarrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle = \hat{c}|\downarrow\downarrow\rangle$ at $N_g = 0$, and $|\uparrow\uparrow\rangle = c\hat{c}|\downarrow\downarrow\rangle$ at $N_g = +1$.

In this formalism, we take the physical states to be the BRST cohomology at ghost number $N_g = 0$. For now we consider states proportional to either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$, and not superpositions of the two.¹¹

In the first case where we take the ghost state to be $|\downarrow\uparrow\rangle$, we find matter/lapse states which are invariant with respect to H and co-invariant with respect to Π . Demanding that $|\psi\rangle \otimes |\downarrow\uparrow\rangle$ is BRST closed requires $H|\psi\rangle = 0$, while $\Pi|\phi\rangle \otimes |\downarrow\uparrow\rangle$ is a BRST exact state $Q(|\phi\rangle \otimes |\downarrow\downarrow\rangle)$ for any matter/lapse state $|\phi\rangle$ satisfying $H|\phi\rangle = 0$. This means that a physical wavefunction $\psi(q, N)$ must obey the Wheeler-DeWitt equation with respect to q (for any fixed N), and among such states any derivative with respect to N is regarded as trivial. This means that the physical state is determined only by the integral $\int_{-\infty}^{\infty} \psi(q, N) dN$, a function depending on q only which solves the Wheeler-DeWitt equation. In particular, we can always choose a representative of the cohomology class which is an eigenstate of N

¹⁰Apart from being conceptually important later for constructing inner products, this also has the technical benefit that the constraint $\Pi = -i\partial_N$ is essentially self-adjoint for real-valued N . If we were to impose the restriction $N > 0$, then Π would have no self-adjoint extension, which later leads to problems with constructing inner products.

¹¹We could more generally look for closed and exact states with a superposition of $|\downarrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$, which might lead to equivalences between the two sectors. This is a little subtle, depending on technical choices regarding which wavefunctions are considered. We will revisit such mixing of the sectors when considering the construction of inner products in section 6.

(with definite lapse $N = N_0$ for any fixed N_0), with wavefunction $\psi(q, N) \propto \delta(N - N_0)$. We can write such a state as $|\psi\rangle \otimes |N = N_0\rangle \otimes |\downarrow\uparrow\rangle$, where $|\psi\rangle$ is an invariant state as discussed above.

Similarly, states proportional to $|\downarrow\uparrow\rangle$ are co-invariant with respect to H , and invariant with respect to Π so the wavefunction $\psi(q, N)$ must be independent of N . The physical states of this form can be written as $|\psi\rangle \otimes |\Pi = 0\rangle \otimes |\downarrow\uparrow\rangle$, where $|\Pi = 0\rangle$ is the state annihilated by Π in the lapse sector (an equal superposition over all N).

Combining these, the cohomology at $N_g = 0$ is a direct sum of invariants and co-invariants (equivalent to (3.6)):

$$\mathcal{H}_{\text{BRST}}^{N_g=0} \simeq V_{\text{inv}} \otimes |N = N_0\rangle \otimes |\downarrow\uparrow\rangle \oplus V_{\text{co}} \otimes |\Pi = 0\rangle \otimes |\downarrow\uparrow\rangle. \quad (3.7)$$

In addition, there is a natural inner product on this $N_g = 0$ cohomology. But once again, this does not give us a norm on invariants or on co-invariants, because the nonzero inner products on the $N_g = 0$ ghost Hilbert space are $\langle \downarrow\uparrow | \downarrow\uparrow \rangle$ and $\langle \downarrow\uparrow | \uparrow\downarrow \rangle$. Instead, we get an anti-symmetric ‘off-diagonal’ pairing between V_{co} and V_{inv} , reproducing the natural pairing we saw above.

3.4 Natural invariant and co-invariant states

We have two different methods of describing physical states. In the next section, we will see how they become equivalent to one another. Nonetheless, some states are simpler or more natural to describe in one language or the other. The summary is that states that involve arbitrary time-evolution in their definition often give rise to invariants (with the evolution projecting onto solutions to the Wheeler-DeWitt equation) [16], while states defined without any evolution (e.g., from ‘initial data’ at a single moment in time) can be interpreted as co-invariants.

3.4.1 Scattering states as invariants

The language of invariants is well-suited to describing states defined asymptotically, in the distant past or future. In the sigma models of section 2.2, the main examples are scattering states in \mathcal{H}_0 : in-states defined at $t \rightarrow -\infty$, or out-states at $t \rightarrow +\infty$.

Concretely, if target space is asymptotically flat in some direction (so $G_{ab}(q)$ and $V(q)$ approach constant values η_{ab} and $\frac{m^2}{2}$ as $q \rightarrow \infty$ in that direction), we can define states

$$|p_{\text{in,out}}\rangle := \lim_{t \rightarrow \pm\infty} e^{-iHt} |p\rangle, \quad (3.8)$$

where $|p\rangle$ is the plane wave wavefunction $e^{ip_a q^a}$ (up to some convenient normalisation, perhaps). The relevant direction taking $q \rightarrow \infty$ is $q^a \propto \pm \eta^{ab} p_b$, with the \pm for out and in states respectively. The momentum should satisfy $\eta^{ab} p_a p_b + m^2 = 0$ to obey the constraint in the ‘free’ region of target space (a more familiar definition of a scattering state would involve an opposite evolution $e^{+iH_0 t}$ with a free Hamiltonian, but we do not need this as H_0 annihilates $|p\rangle$). With this condition, we generally expect (3.8) to approach a nonzero limiting wavefunction which satisfies the constraints.

For an alternative definition of scattering states (which generalises to target spaces with different asymptotics), we can work more directly in a position basis. Intuitively, we define a state that was at some definite point in target space q in the past (for an in-state), and take q to infinity in some direction. To make this precise, define

$$|p_{\text{in,out}}\rangle := \lim_{q \rightarrow \infty} \mathcal{N}(q) \int_{t_0}^{\pm\infty} dt e^{-iHt} |q\rangle, \quad (3.9)$$

where $\mathcal{N}(q)$ (usually a phase) is chosen so that the limit exists and is nonzero and t_0 is arbitrary since the result comes from large $|t|$ in the limit. Choosing $t_0 = 0$, we can write the integral as an operator $\frac{1}{iH \pm \epsilon}$ acting on $|q\rangle$; this gives a Green's function which does not solve the constraints for finite q , but instead gives a delta-function source on the right hand side. But when we take $q \rightarrow \infty$, this source gets pushed off to infinity and so the resulting state solves the Wheeler-DeWitt equation with no source.

In the most physically important examples of such scattering states, the relevant $q \rightarrow \infty$ region of superspace corresponds to a large spatial volume, with the direction specifying the conformal geometry and the behaviour of matter fields. An example is defining states in asymptotically de Sitter universes by data at future/past infinity \mathcal{I}_{\pm} (see [5] for an example in de Sitter JT gravity).

Note that there may be some direction q^0 in target space which is natural to call ‘target space time’ (particularly if G_{ab} is Lorentzian). This need not be correlated with our ‘worldline’ time. In particular, for an asymptotically Minkowski target space (with $V(q) \sim \frac{m^2}{2} > 0$ as $q^0 \rightarrow \pm\infty$) we expect four classes of scattering states: in- and out-states, with p_0 positive or negative. From the perspective of worldline QFT, these four choices correspond to solutions to the equations of motion which are purely positive or negative frequency in the asymptotic past or future of target space. These will not be independent states; the ‘S-matrix’ of the worldline quantum mechanics relating in- and out-states describes the Bogoliubov coefficients of the QFT. We discuss these states explicitly in section 4.5.

3.4.2 Finite time states as co-invariants

While invariants are a useful description of asymptotic states, co-invariants give us a natural way to define states at some finite time. For this, we simply choose some arbitrary wavefunction $|\psi\rangle \in \mathcal{H}_0$ in the unconstrained Hilbert space as a representative of the coset $|\psi\rangle$. In particular, we can choose states corresponding to a definite $q = q_0$ (e.g., fixed spatial metric in a minisuperspace model), with wavefunction $\psi(q) = \delta(q - q_0)$. Other useful choices include coherent states which are well-localised in both position and momentum space (e.g., both metric and extrinsic curvature) in the classical limit.

This is somewhat analogous to defining a classical spacetime metric by initial data at one time. The advantage is that we can describe physics in a way which is local in time. The price we pay is that the same state may be described in many different ways: for any t , $e^{-iHt}|\psi\rangle$ describes the same coset $|\psi\rangle$ defined at a different initial time (and we can take any superposition over such time translations). We comment more on the connection between this and gauge fixing in the next subsection.

3.5 States in the path integral formalism

In the path integral formalism, states are associated with choices of initial or final boundary condition. The different classes of states we have discussed above both appear naturally in the path integral formalism, but in slightly different ways. Understanding this is very useful for giving a proper Hilbert space interpretation to various gravitational path integrals.

First, the co-invariant states (particularly those with delta-function representative wavefunctions $\psi(q) = \delta(q - q_0)$ as above) appear very naturally in the path integral as boundary conditions at a fixed finite time. For example, we might be studying a path integral in pure gravity with a Dirichlet boundary condition of fixed spatial metric g at initial time $t = 0$, and (for a properly gauge-invariant path integral) we can interpret this boundary condition as a co-invariant state $|g\rangle$. Given the redundancy inherent in such states, it is natural to ‘gauge fix’ by restricting to some class of boundary conditions, which we discuss in a moment.

In contrast, the invariant scattering states introduced above can be defined in the path integral by boundary conditions in the infinite asymptotic past or future. This is manifest from (3.8), for example, which corresponds to Neumann boundary conditions fixing \dot{q} (in a configuration space path integral) or p (in a first-order phase space path integral) at $t \rightarrow \pm\infty$.

Given this, it is straightforward to define a path integral which computes the pairing $\langle\langle \cdot | \cdot \rangle\rangle$ between co-invariant and invariant (scattering) states: we simply integrate over all fields on a semi-infinite half-line (say, $t < 0$). We can remove the diffeomorphism invariance by applying a local gauge-fixing condition at each point in the bulk, most simply fixing lapse $N = 1$ so that coordinate time coincides with proper time. This does not leave any residual gauge symmetries: rigid time translation is allowed locally, but disallowed by fixing the special $t = 0$ boundary. Nor does it leave any moduli to integrate over (since there is a unique metric on the half-line up to diffeomorphisms). These facts mean that we do not have to make any further choices to define the path integral computing something like $\langle\langle q | \psi_{\text{in}} \rangle\rangle$. This straightforwardness is a reflection of the fact that this pairing between invariants and co-invariants is canonical. Finally, note that $\langle\langle q | \psi_{\text{in}} \rangle\rangle$ can also be interpreted as the wavefunction $\psi_{\text{in}}(q)$ of the invariant state (in the same way that $\langle q | \psi \rangle$ is the position-space wavefunction in ordinary quantum mechanics), which will satisfy the Wheeler-DeWitt equation as a concrete differential equation (rather than a more abstract operator version).

These boundary conditions also naturally give the correct states on the ghosts in the BRST formalism. The Fadeev-Popov ghost field $c(t)$ (associated with the Hamiltonian constraints) is a fermionic counterpart to the time component of vector fields $\xi(t)\partial_t$ which generate diffeomorphisms. For boundary conditions defined at a fixed finite time $t = 0$, these diffeomorphisms must vanish on the boundary, $\xi(t = 0) = 0$. This Dirichlet boundary condition is inherited by the ghosts giving $c(t = 0) = 0$, which in the BRST Hilbert space results in states satisfying $c|\psi\rangle = 0$. This defines the ‘top’ level of cohomology, where the co-invariants live. There is no such constraint for the scattering states, leading to the opposite ‘Neumann’ boundary condition $b = 0$.

3.6 Co-invariants and gauge-fixing

While co-invariant states can be useful for making locality in time manifest, we must deal with the enormous ambiguity from the equivalence under acting with constraints. But such ambiguity is not unfamiliar: it is the usual redundancy from gauge-invariance, which we typically can deal with by choosing a gauge.

Quantum mechanically, gauge-fixing a state means we would like to specify some way to choose a unique representative from each co-invariant coset. For example, in a sigma model we can divide our configuration space coordinates into a ‘clock’ χ and transverse coordinates q_\perp , and specify the state at a time when $\chi = 0$. Classically, this means specifying an $H = 0$ trajectory in phase space by a point at which it intersects the surface Σ_χ defined by $\chi = 0$: that is, by initial data satisfying the gauge condition $\chi = 0$. A quantum version of this is to make an ansatz (with two terms because H is quadratic in momenta, as explained below)

$$\psi(q) = \|d\chi\| \delta(\chi) \psi_1(q_\perp) + \delta'(\chi) \psi_2(q_\perp). \quad (3.10)$$

The factor $\|d\chi\|$ of the norm of the derivative of χ (under the target space metric) is chosen to make this invariant under reparameterisations of the clock variable. An analogue in a higher-dimensional theory of gravity is to choose a function χ of spatial metric and matter fields at each position in space, so that (classically) $\chi = 0$ defines a Cauchy surface for the spacetimes of interest.

Ideally, each coset in V_{co} will have a unique wavefunction representative of this form: that is, the map from a pair of functions (ψ_1, ψ_2) of q_\perp to V_{co} defined by (3.10) is bijective. This could fail in two ways. Either some cosets do not have a representative (the map is not surjective), which means that this ansatz will not describe the whole physical Hilbert space; the classical analogue is a trajectory that never encounters the surface Σ_χ in phase space. Alternatively, there may be several states that give rise to the same coset (so that there is some nonzero ‘pure gauge’ (ψ_1, ψ_2) in the image of H). In that case we have not completely eliminated the redundancy in the co-invariant description; the classical analogue is a trajectory which crosses Σ_χ several times.

Our ansatz is chosen to satisfy this existence and uniqueness locally at $\chi = 0$ and perturbatively in diffeomorphisms, by which we mean that there is a unique representative among wavefunctions supported on the $\chi = 0$ surface, imposing triviality of the image of H among such states (which is weaker than using finite transformations e^{-iHt}). We included terms proportional to $\delta(\chi)$ and $\delta'(\chi)$ since H is quadratic in momenta (a quantum analogue of initial data requiring both position and velocity), allowing us to remove second and higher derivatives of $\delta(\chi)$. It is reasonable to expect that a sufficiently smooth wavepacket supported in a neighborhood of $\chi = 0$ will also have a representative of the form (3.10), at least approximately (for example, one strategy is to Fourier transform from χ to the conjugate momentum space variable p_χ , split the resulting wavefunction ψ into even and odd parts in p_χ , and eliminate p_χ^2 using the constraint).¹²

¹²We can investigate whether $\chi = 0$ provides a ‘good gauge’ non-perturbatively via the pairing with invariants, checking that the ansatz (3.10) preserves non-degeneracy of the form $\llbracket \phi | \psi \rrbracket = \int \phi^* \psi$. That is, the pairing should vanish for all invariants $|\phi\rangle$ only if $\psi_1 = \psi_2 = 0$ (otherwise there are residual gauge

For applications to more realistic theories of gravity, we propose this as a practical way to describe the physical Hilbert space by local data, working in perturbation theory around some classical background. This means making a gauge-fixed ansatz analogous to (3.10), where the choice corresponds to a gauge which is good classically for spacetimes sufficiently close to the background. For this, it is probably most convenient to replace the delta functions with narrow Gaussians, so that the states correspond to classical field configurations (coherent states, well-localised in position and momentum). We will later give a systematic procedure to construct such an ansatz from a classical gauge choice (including Gaussian averaging). We then expect the gauge-fixed ansatz to take care of all redundancy in perturbation theory. Nonetheless, there can still be non-perturbative redundancies; we propose these as a possible gauge symmetry explanation for the existence of null states in effective gravity. We revisit this point in the discussion section 8.

3.6.1 Maximal volume gauge

Of course, there are many possible choices of gauge condition $\chi = 0$, where $\chi(q, p)$ can depend on momenta as well as positions. Each gauge condition should lead to a different possible ansatz like (3.10). An important example for higher-dimensional theories of gravity is the ‘maximal volume’ (or zero mean curvature) gauge choice [15]. For this, write the spatial metric $g = \Omega^2 \hat{g}$ in terms of a conformal metric \hat{g} (fixed to have unit determinant, for example) and conformal factor Ω , and choose a wavefunction to be independent of Ω . This means that the state is a delta function in the canonical conjugate to Ω , which is proportional to the mean curvature (trace of extrinsic curvature) $\text{Tr } K$. An analogue for a one-dimensional sigma model (2.4) is to choose a gauge which is linear in momentum, such as $\chi = p_0$ where p_0 is the momentum conjugate to some target space coordinate q^0 . In a mini-superspace model, q^0 might be interpreted as parameterising the volume of space (in which case it is often the timelike direction in a Minkowski-signature target space). This is a good gauge when classical trajectories involve q_0 increasing to a unique maximum before decreasing. This could model a big bang/big crunch cosmology, or alternatively a family of Cauchy surfaces in an asymptotically AdS universe all sharing the same asymptotics at spatial infinity. In two-dimensional JT gravity (either for closed universes or a two-sided Hilbert space with asymptotically AdS boundary conditions), a similar interpretation applies to the geodesic states studied in [30, 31], with χ being the momentum conjugate to the dilaton (which is proportional to extrinsic curvature). In our language these can be

equivalences), and it should vanish for all $\psi_{1,2}$ only if $[\phi] = 0$ (otherwise, there are co-invariants without a representative of this form). The pairing results in an integral over the surface Σ_c (here meaning a surface in configuration space \mathcal{M} rather than phase space $T^*\mathcal{M}$):

$$[\phi|\psi\rangle\rangle = \int_{\Sigma_\chi} dq_\perp [\phi^* \psi_1 - (\nabla_N \phi^*) \psi_2]. \quad (3.11)$$

The measure is the volume form on Σ_χ induced from the target space metric G , and ∇_N the normal derivative. Resemblance to the Klein-Gordon form (which we will meet below) is not a coincidence. The nondegeneracy of (3.11) is equivalent to existence and uniqueness of solutions to $H[\phi] = 0$ given initial data $(\phi, \nabla_N \phi)$ on Σ_χ . This holds if \mathcal{M} is Lorentzian and globally hyperbolic and Σ_χ is a Cauchy surface [29], so in such cases (3.10) is a good gauge (but is likely to fail otherwise).

interpreted as co-invariants with wavefunctions chosen to be independent of the dilaton (or a delta function in its canonical conjugate, the extrinsic curvature).

4 Gravitational inner products

From the above discussion of spaces of states, two issues remain to be resolved in order to define a physical Hilbert space of the constrained theory. The first is that we apparently have two different ways to define a space of states, so we must decide whether to use invariants or co-invariants. The second is that we do not immediately have an inner product on either of these spaces of states.

In this section we point out that both of these are resolved together: an inner product on either V_{co} or V_{inv} gives us an equivalence between these two spaces of states. We then discuss specific proposals for the inner product. First we describe the ‘group average’ inner product and related ideas, which we believe to be the most natural construction for the gravitational Hilbert space. We then discuss the Klein-Gordon form for sigma models, which was the originally proposed inner product of DeWitt [7]. We explain how each of these arise in the path integral formulation, in particular how the Klein-Gordon form is obtained from gauge fixing [32] and its relation to the gauge-fixed co-invariants in section 3.6. We then illustrate these ideas explicitly in models with one-dimensional target spaces. Finally, we contrast these constructions with the Hilbert space obtained in the worldline formulation of QFT.

4.1 Inner products and equivalences between invariants and co-invariants

The spaces of co-invariants and invariants are not unrelated: as noted above they are dual spaces, with the natural nondegenerate sesquilinear pairing $\langle\langle \cdot | \cdot \rangle\rangle$ between V_{inv} and V_{co} (and hence the anti-linear bijection $V_{\text{inv}} \simeq V_{\text{co}}^*$). With this, a Hermitian form on V_{co} is equivalent to a linear Hermitian map $\eta : V_{\text{co}} \rightarrow V_{\text{inv}}$ (usually called a *rigging map*) by defining

$$\langle\langle \psi_2 | \psi_1 \rangle\rangle_\eta := \langle\langle \psi_2 | \eta \psi_1 \rangle\rangle = \llbracket \eta \psi_2 | \psi_1 \rrbracket, \quad |\psi_{1,2}\rangle \in V_{\text{co}}. \quad (4.1)$$

The second equality is a requirement on η to ensure Hermiticity of $\langle\langle \cdot | \cdot \rangle\rangle_\eta$. It is equivalent to $\eta^\dagger = \eta$, with \dagger defined with respect to the pairing $\llbracket \cdot | \cdot \rrbracket$ (so $\eta^\dagger : V_{\text{co}} \rightarrow V_{\text{inv}}$ is defined by $\langle\langle \psi_2 | \eta^\dagger \psi_1 \rangle\rangle = \llbracket \eta \psi_2 | \psi_1 \rrbracket$).

If this form is positive definite we have an inner product on co-invariants, and passing to the completion with respect to the resulting norm on V_{co} gives us a Hilbert space \mathcal{H}_{co} of co-invariants.¹³ We will see interesting examples below which do not obey this, but only satisfy a weaker condition that the form is non-degenerate (i.e., if $\langle\langle \psi_2 | \psi_1 \rangle\rangle_\eta = 0$ for all $|\psi_2\rangle$, then $|\psi_1\rangle = 0$), which is equivalent to η being injective ($\llbracket \eta \psi \rrbracket = 0 \implies |\psi\rangle = 0$). This means that η is a linear isomorphism between co-invariants V_{co} and a subspace $\text{im } \eta \subseteq V_{\text{inv}}$

¹³To get a good Hilbert space it is sufficient for $\langle\langle \cdot | \cdot \rangle\rangle_\eta$ to be positive *semi*-definite, in which case the resulting Hilbert space can informally be thought of as the quotient $\mathcal{H}_{\text{co}} \simeq V_{\text{co}} / \ker \eta$ of all co-invariants by the ‘null states’ $\ker \eta$ (in addition to the null states $\text{im } H \subseteq \mathcal{H}_0$ already implicit in the definition of V_{co}). But we see no reason to expect additional null states, unless there is some additional (perhaps discrete) gauge symmetry which could be accounted for in the original definition of V_{co} .

of invariants. It is reasonable that this subspace might not be the whole of V_{inv} , since we would like to exclude invariant wavefunctions corresponding to non-normalisable states, for example (comments on examples are given in footnote 17). The upshot is that putting an inner product on co-invariants gives us an equivalence between our two ways of describing states, and vice-versa.

It is slightly more concrete to think of η as an operator on the original Hilbert space \mathcal{H}_0 ,¹⁴ and this is the perspective we will take for the most part. To get this from the above map $V_{\text{co}} \rightarrow V_{\text{inv}}$ we simply map any state $|\psi\rangle \in \mathcal{H}_0$ to its co-invariant coset $|\psi\rangle\rangle$ before applying η , and regard the resulting invariant $|\eta\psi\rangle$ as a (distributional) state in \mathcal{H}_0 . In this way of thinking we require $\eta H = 0$ (because η is well defined on co-invariant cosets) and $H\eta = 0$ (because η maps to invariants). It will also be Hermitian $\eta^\dagger = \eta$ with respect to the usual \mathcal{H}_0 inner product (which makes the two properties above equivalent), and positive semi-definite to get a good Hilbert space (with $\text{im } H$ giving null states).

The map η also induces an equivalent Hilbert space structure on invariants. Assuming η is nondegenerate as above, it has an inverse $\eta^{-1} : \text{im } \eta \subseteq V_{\text{inv}} \rightarrow V_{\text{co}}$. we get a Hermitian form on invariants by inverting η :

$$\llbracket \psi_2 | \psi_1 \rrbracket_\eta = \llbracket \psi_2 | \eta^{-1} \psi_1 \rrbracket = \langle \langle \eta^{-1} \psi_2 | \psi_1 \rangle \rangle = \langle \langle \eta^{-1} \psi_2 | \eta^{-1} \psi_1 \rangle \rangle_\eta, \quad |\psi_{1,2}\rangle \in \text{im } \eta. \quad (4.2)$$

As for co-invariants, if η is positive-definite this is an inner product, and by taking the completion we turn $\text{im } \eta \subseteq V_{\text{inv}}$ into a Hilbert space \mathcal{H}_{inv} , which is naturally isomorphic to \mathcal{H}_{co} using η .

Once again, we would like to make this more concrete and practical by thinking about η^{-1} as a map on the original Hilbert space \mathcal{H}_0 . This means we want an operator which takes a solution $|\psi\rangle$ to the Wheeler-DeWitt equation $H|\psi\rangle$, and gives us some state $|\phi\rangle$ such that $\eta|\phi\rangle = |\psi\rangle$. But this leaves us with a lot of freedom: we can choose any gauge-equivalent representative $|\phi\rangle$ from its co-invariant coset. As a result, η^{-1} is represented concretely by a whole family of maps, any one of which we call κ . The condition for κ to be compatible with η is that it is a generalised inverse, meaning

$$\eta\kappa\eta = \eta. \quad (4.3)$$

One way to read this equation is that κ is a right inverse to η (i.e., $\eta\kappa$ is the identity) when acting on the image of η (in particular, on invariant wavefunctions). Another way to read it is that κ is a left inverse to η , up to addition of a null state in $\ker \eta$.

Since κ selects some particular representative from a co-invariant coset, we can think of it as a *gauge-fixing map*. For example, following the discussion of section 3.6 we might like to choose κ to fix to the $\chi = 0$ gauge, which means that it always maps invariants to wavefunctions of the form (3.10). Such a map is nice for recovering a description which is local in time (on a particular gauge-fixed Cauchy surface) from a Wheeler-DeWitt wavefunction which is highly non-local in time, involving a superposition over all possible Cauchy surfaces. We will see later how to construct such a κ systematically from any gauge condition.

¹⁴More precisely, it's a distributional operator which maps test function states Φ to distributional states in $\bar{\Phi}^*$, in the notation of section 3.2.

The challenge (arising from Gribov ambiguities, and the impossibility of finding a universally applicable gauge condition) is to perform such a gauge-fixing while simultaneously satisfying the compatibility condition (4.3).

The various maps and inner products introduced in this section and the relations between them were sketched in figure 1.

This discussion is rather abstract, so we now turn to examples which illustrate these ideas. At this point there is a great deal of freedom in the selection of an inner products (η or κ) with the above properties, so these examples may seem like arbitrary choices. We will give a more principled justification these choices later when we discuss how these inner products arise in the BRST formalism, which provides a much more rigid framework.

4.2 Group averaging inner product

There is one choice of a rigging map which does not depend on any details of the unconstrained theory besides the abstract Hamiltonian constraint H , and automatically gives us a positive inner product. It also has a simple, natural and general path integral description (as we will describe in section 4.4), and also emerges naturally in the BRST formalism (section 6). For these reasons, we advocate that this is the leading candidate for the physical inner product.

As a map on \mathcal{H}_0 , we can write this formally as a delta function of the constraint:

$$\eta = 2\pi\delta(H). \quad (4.4)$$

This is manifestly Hermitian, positive semi-definite (since δ is a non-negative distribution), and satisfies $\eta H = H\eta = 0$ as required for a rigging map.

Another suggestive way to write this is as a ‘group average’, integrating over the action of the gauge group of time translations:

$$2\pi\delta(H) = \int_{-\infty}^{\infty} dt e^{-iHt}. \quad (4.5)$$

This integral will converge when we take the matrix elements of states with sufficiently smooth energy resolution. This construction generalises rather nicely to many other groups G of gauge transformations. For that, take an unconstrained Hilbert space \mathcal{H}_0 which furnishes a unitary representation $U(g)$ of G ; then η is constructed by integrating $U(g)$ over G with an appropriate measure [33]. In the case that G is compact (so that it has a discrete set of irreducible unitary representations), η is simply the orthogonal projection onto the trivial representation, the subspace $V_{\text{inv}} \subseteq \mathcal{H}_0$ of gauge-invariant states. But the construction applies also for many non-compact groups, such as the case \mathbb{R} of interest here, as well as more complicated situations encountered in gravity including constraint algebras with field-dependent structure constants [20].

Note that this generalisation also naturally incorporates global information about the gauge group; this is in contrast to BRST constructions which depend only on the Lie algebra of infinitesimal gauge transformations.

This η can be thought of either as a map $V_{\text{co}} \rightarrow V_{\text{inv}}$, which ‘projects’ a general state onto a solution of the Wheeler-DeWitt equation, or as an inner product on V_{co} .

As explained above, this η also induces an equivalent dual inner product κ on invariants. To understand this informally, recall that energy eigenstates in \mathcal{H}_0 are delta-function normalised, with $\langle \psi_{E'} | \psi_E \rangle \propto \delta(E - E')$. In particular, invariants (zero energy eigenstates) have an infinite norm due to a factor of $\delta(E = 0)$. The κ dual to the group-average η gives a finite inner product by dividing out by this infinite factor (times 2π).

The ‘spectral analysis’ inner product makes this mathematically precise [34]. We use the spectral theorem for a self-adjoint constraint H , which tells us that the Hilbert space decomposes as a direct integral of energy eigenspaces \mathcal{H}_E .¹⁵

$$\mathcal{H}_0 = \int^{\oplus} dE \mathcal{H}_E. \quad (4.6)$$

This means that a state $|\psi\rangle \in \mathcal{H}_0$ is given by specifying a family of states $|\psi_E\rangle \in \mathcal{H}_E$ for each energy E , and the inner product is computed by the integral $\langle \psi' | \psi \rangle = \int \frac{dE}{2\pi} \langle \psi'(E) | \psi(E) \rangle$. We can then simply define the invariants as the zero energy eigenspace, $V_{\text{inv}} := \mathcal{H}_{E=0}$. This space already comes equipped with an inner product, and this is precisely the κ described informally above. In this language, $\eta = 2\pi\delta(H)$ is the map taking the family $\{|\psi(E)\rangle\}_E$ onto its zero energy component $|\psi(E=0)\rangle$. A generalised inverse κ takes a zero-energy state $|\psi(E=0)\rangle$, and extends it in an arbitrary way to a smooth family $\{|\psi(E)\rangle\}_E$.

However, this spectral analysis construction of the inverse only directly generalises for abelian constraints. One of the challenges for practically applying the group averaging procedure is that there is not more generally an obvious recipe to write down or systematically compute the dual inner product κ on invariant states. This is one motivation for exploring other possibilities and understanding how they relate to group averaging.

4.3 The Klein-Gordon form

While there are many reasons to advocate for the group averaging inner product, it was not the first proposal for gravity. This was the ‘Klein-Gordon’ form on the space of invariants suggested by DeWitt in analogy to QFT [7]. This has the drawback that it is not positive-definite, so it is not truly an inner product (we will discuss the relation to the inner product in free QFT in section 4.6). But while it has indefinite signature, it is non-degenerate so most of the above formalism continues to apply. Some useful references in the context of QFT are [35, 36].

For this construction, we need more structure than the abstract Hamiltonian H . We specialise to the sigma-model examples of section 2.2, for which the Hamiltonian is given by $H = -\frac{1}{2}\nabla^2 + V(q)$, where ∇^2 is the Laplacian on a pseudo-Riemannian manifold \mathcal{M} and V a potential function on that target space. In fact, we will assume that the target space metric G_{ab} is Lorentzian and such that \mathcal{M} is globally hyperbolic. This means that the constraint equation $H[\psi] = 0$ has a good Cauchy problem. Specifically, these conditions guarantee that solutions $\psi(q)$ are in one-to-one correspondence with initial data on a Cauchy surface Σ (the value $\Phi = \psi|_{\Sigma}$ and normal derivative $\Pi = \nabla_N \psi|_{\Sigma}$ of the field ψ). For a short review

¹⁵Here dE is the usual flat measure on \mathbb{R} (up to an additional factor of 2π appearing in the inner product), and we have assumed that H has a purely continuous spectrum. The construction goes through unaltered if H has additional discrete spectrum, as long as there is a neighborhood of $E = 0$ with no discrete eigenvalues.

of relevant technical results, see [29]. We emphasise that this is an initial value problem with respect to ‘target space time’, not to be confused with the ‘worldline’ time t which is the physical time from the perspective of our one-dimensional theory of gravity.

In this specific context, the invariant states are precisely the solutions to the Wheeler-DeWitt equation $H|\psi\rangle = 0$. More precisely, we may take V_{inv} to be the space of complex solutions $\psi(q)$ with smooth compactly supported initial data (Φ, Π) on some Cauchy surface Σ (a condition which is independent of Σ). On this space of invariants define the ‘symplectic bilinear form’

$$\Omega(\psi_2, \psi_1) = \frac{1}{2} \int_{\Sigma} (\Pi_2 \Phi_1 - \Phi_2 \Pi_1), \quad (4.7)$$

where Σ is any Cauchy surface (we always take the measure of integrals on Σ as the volume form induced from the target space metric G). This is independent of the choice of Σ . Note that the definition requires us to choose a time-orientation for \mathcal{M} , since the initial velocity $\Pi = \nabla_N \phi$ is the normal derivative of the field with future-pointing unit normal vector N . By combining with the ‘complex structure’ $\psi \mapsto \bar{\psi}$ we define the Klein-Gordon form

$$\llbracket \psi_2 | \psi_1 \rrbracket_{\text{KG}} := i \Omega(\bar{\psi}_2, \psi_1), \quad (4.8)$$

which is a non-degenerate (but not positive) Hermitian form on V_{inv} . The constant of proportionality must be imaginary for Hermiticity, but is otherwise a matter of convention.

Following the earlier general discussion, we can also characterise the Klein-Gordon form by relating it to structures involving the dual space V_{co} . If V_{inv} is the space of solutions to $H\psi = 0$, we can think of V_{co} as the space of *sources* f for the inhomogeneous equation $H\psi = f$. Take co-invariants $|f\rangle\rangle$ to be represented by functions (or more generally distributions) f on target space \mathcal{M} with the natural pairing $\langle\langle f | \psi \rangle\rangle = \int_{\mathcal{M}} \bar{f} \psi$. For this to be well-defined, we take f to have compact support in time,¹⁶ and to define V_{co} we quotient by the image of H in this space (since $\langle\langle Hg | \psi \rangle\rangle = 0$ for any solution ψ). The compactness of f in time is a very important condition, since without it any function would be in the image of H and the space of co-invariants would become trivial. Now, the abstract results above tells us that the Klein-Gordon form is equivalent to an isomorphism between co-invariants and invariants (maps $\eta : V_{\text{co}} \rightarrow V_{\text{inv}}$ and $\eta^{-1} = \kappa : V_{\text{inv}} \rightarrow V_{\text{co}}$), and to a Hermitian form $\langle\langle \cdot | \cdot \rangle\rangle_{\text{KG}}$ on co-invariants.

To make this concrete, consider the ‘causal propagator’ (or Pauli-Jordan distribution) Δ which constructs solutions $\psi = \Delta f$ from a source f . This solution is defined as the difference between retarded $\mathcal{G}_+ f$ and advanced $\mathcal{G}_- f$ solutions to the constraint equation with source f . In other words,

$$\begin{aligned} \Delta f &= \mathcal{G}_+ f - \mathcal{G}_- f, \\ H\mathcal{G}_{\pm} f &= f, \quad \mathcal{G}_{\pm} f = 0 \text{ to the past(+)/future(-) of the support of } f. \end{aligned} \quad (4.9)$$

¹⁶This means that the support of f must be bounded by a pair of Cauchy surfaces. We could be less strict and allow non-compact support with sufficiently fast decay in the future, but this wouldn’t add anything since every coset in fact has a representative with compact support in time.

This Δ gives us a linear map $V_{\text{co}} \rightarrow V_{\text{inv}}$ from co-invariants to invariants, since Δf solves the constraint equation, and if f is in the image of H then $\Delta f = 0$. For the latter, note that if $f = Hg$ for some compactly supported g then g serves as both the advanced and retarded solution with source f , so $\mathcal{G}_+ f = \mathcal{G}_- f = g$. It is also injective: suppose $\Delta f = 0$ and let $g = \mathcal{G}_+ f = \mathcal{G}_- f$; then g is compactly supported and $f = Hg$, so $f \in \text{im } H$. Surjectivity will be proved by explicitly constructing an inverse in a moment, so we have an isomorphism.¹⁷

The causal propagator is related to the Klein-Gordon form by the identity

$$\int_{\mathcal{M}} \psi f = \Omega(\psi, \Delta f), \quad (4.10)$$

which we prove in appendix A. In terms of our pairing between invariants and co-invariants we can rewrite (4.10) as

$$\llbracket \psi | f \rrbracket = -i \llbracket \psi | \Delta f \rrbracket_{\text{KG}} \implies \eta_{\text{KG}} | f \rrbracket = -i | \Delta f \rrbracket, \quad (4.11)$$

where in the second line we have identified the map $\eta_{\text{KG}} : V_{\text{co}} \rightarrow V_{\text{inv}}$ associated with the Klein-Gordon form as $-i\Delta$. This immediately gives us a Hermitian form on co-invariants, which we can write very concretely in terms of the kernel $\Delta(q_2, q_1)$ of the causal propagator,

$$\llbracket f_2 | f_1 \rrbracket_{\text{KG}} = -i \int_{\mathcal{M}} \bar{f}_2(q_2) \Delta(q_2, q_1) f_1(q_1) dq_1 dq_2. \quad (4.12)$$

Note that the group average inner product can also be written as a difference between two propagators [27]:

$$\eta_{GA} = 2\pi\delta(H) = i \left(\frac{1}{H + i\epsilon} - \frac{1}{H - i\epsilon} \right). \quad (4.13)$$

In term of the group average integral (4.5), the terms come from positive and negative t respectively.¹⁸ This expression is ‘Feynman minus anti-Feynman’ rather than ‘advanced minus retarded’. We can think of this as coming from a fixed-energy causal (retarded minus advanced) propagator for H with respect to *worldline time* (while the Klein-Gordon norm comes from the causal propagator with respect to target space). This hints at a relationship between the Klein-Gordon form and group averaging, which we will see in various guises later: they agree when the proper (worldline) time and target space time flow in the same direction, and disagree by a sign when they flow in opposite directions.

Finally, we can also construct the corresponding map κ from invariants to co-invariants. Concretely, this means finding a family of maps taking a solution ψ to a source f satisfying

¹⁷Follow the construction as outlined in section 3.2, we might choose to start with test functions Φ being smooth and compactly supported (or rapidly decaying) wavefunctions, with a dual space Φ^* of distributions (or tempered distributions) on \mathcal{M} . Then $\text{im } \eta_{\text{KG}}$ will consist of solutions with smooth and compactly supported (or rapidly decaying) initial data. This would be a strict subspace of V_{inv} , which would contain all (tempered) distributional solutions with no requirements on smoothness or decay of initial data. For many of the states in V_{inv} but not in $\text{im } \eta_{\text{KG}}$, the Klein-Gordon form will not converge so it is necessary to exclude them. Similar comments apply to group averaging, except that this does not respect target space lightcones so $\text{im } \eta$ is not restricted to compactly supported initial data.

¹⁸This split is rather special to one dimension, where evolution is either forward or backward in proper time. In higher dimensions you also get contributions from a mixture, going forward in some regions of space and backwards in others.

$\psi = -i\Delta f$ (with any two such sources differing by Hg for some compactly supported g). Since the Klein-Gordon form can be expressed as an integral on a Cauchy surface Σ , it is natural to guess that we should be able to find a source f for any ψ with compact support contained in Σ . Assuming for now that such a source exists, restricting the solution $\psi = -i\Delta f$ to the future of Σ would give us just the retarded solution: defining a step function $\Theta_\Sigma(q)$ equal to 1 to the future of Σ and zero to the past, we have $\Theta_\Sigma(q)\psi(q) = -i\mathcal{G}_+ f(q)$ (plus a compactly supported distribution g on Σ , perhaps). Then we can recover f (plus $Hg \in \text{im } H$ which leads to the same coset in V_{co}) by acting with the Klein-Gordon operator H :

$$f_\Sigma = iH(\Theta_\Sigma\psi). \quad (4.14)$$

This explicitly constructs the desired compactly supported distribution inverting $\eta|f\rangle = |\psi\rangle$, so $\kappa|\psi\rangle = |f_\Sigma\rangle$. Despite appearances, f_Σ is not in the image of H because $\Theta_\Sigma\psi$ is not compactly supported in the future. More generally, we can replace Θ_Σ with any function equal to one in the future and zero in the past (or any other constant values differing by unity). We can also use this result to write the Klein-Gordon form in an alternative way,

$$\kappa_\Sigma = i[H, \Theta_\Sigma], \quad (4.15)$$

where we write a commutator so that κ_Σ is a more manifestly Hermitian operator (though this still needs care, since H is not Hermitian on invariants due to boundary terms when integrating by parts). To recover (4.7), note first that integrals on Σ can be rewritten in term of integrals on \mathcal{M} (with appropriate measures) using the step function Θ_Σ :

$$\int_\Sigma N_a V^a = \int_{\mathcal{M}} V^a \nabla_a \Theta_\Sigma. \quad (4.16)$$

We have $[H, \Theta_\Sigma] = -\frac{1}{2}(\nabla^a \nabla_a \Theta_\Sigma + \nabla_a \Theta_\Sigma \nabla^a)$, so the matrix elements of this commutator (after integrating by parts so the first term acts to the left) indeed give us (4.8).

If we choose a ‘clock function’ χ as one of our coordinates as in section 3.6 so that the Cauchy surface Σ corresponds to $\chi = 0$ and $\Theta_\Sigma(q) = \Theta(\chi)$, then the wavefunction (4.14) resulting from the κ map will precisely take the form of the ansatz (3.10) we gave for our gauge-fixed co-invariant representative. Concretely, this arises from the product rule for the Laplacian in H where either one derivative or two hit the factor of $\Theta(\chi)$, leading to $\delta(\chi)$ or $\delta'(\chi)$ terms (and terms where Θ is not differentiated do not contribute, using $H\psi = 0$). This suggests a general way to construct such a gauge-fixing ansatz: we require that the wavefunctions are in the image of some particular κ map. Later we will see how to construct much more general classes of κ map systematically from any gauge choice, allowing us to write ansätze analogous to (3.10) for any gauge.

4.4 Path integrals, inner products and gauge fixing

Several of the constructions we have mentioned here arise naturally in the path integral formulation. To understand the various possibilities, we first recall from section 3.5 how co-invariant and invariant states often arise as boundary conditions in the path integral: co-invariants often are associated with boundaries at a finite time, while invariants often arise

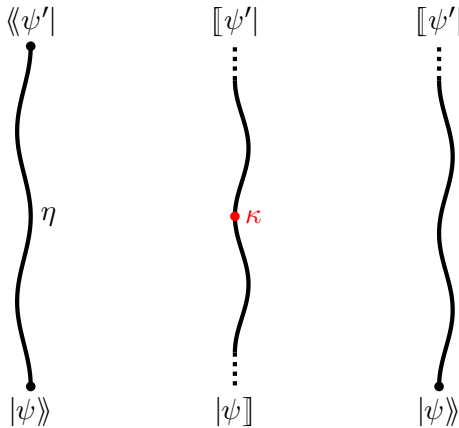


Figure 3. Pairings between various types of states in one-dimensional gravity arise naturally from path integrals over different sorts of intervals. A path integral over a finite interval (left) gives a pairing between co-invariants; fixing to constant lapse gauge leaves a single modulus, and integrating this over \mathbb{R} gives the group-average η . A path integral over an interval which is infinite at both ends (middle) naturally describes a pairing between invariants (described by asymptotic boundary conditions); there is a residual time-translation symmetry which must be gauge-fixed by insertion of a κ map. Finally, a semi-infinite integral has neither a modulus nor a residual symmetry, and it leads to the simple canonical pairing between one invariant and one co-invariant state.

from boundary conditions in the infinite past or future. When we try to pair these with one another, for one-dimensional gravity we arrive at spacetimes which look like various types of intervals. From a pairing of two co-invariants (where we expect an η map) we get a finite interval $[T_i, T_f]$; from two invariants (expecting a κ map) we get the whole real line \mathbb{R} ; if we have one of each we get a semi-infinite interval $[T_i, \infty)$ or $(-\infty, T_f]$. We already saw in 3.5 that the last option naturally and immediately gives us the canonical pairing between the two types of state $[[\cdot]\rangle\rangle$, but for the other two possibilities different issues arise. Before we look at each of these, string theorists may like to note an analogy with various possible worldsheet topologies: for the propagation of closed strings these intervals are analogous to a cylinder, a sphere two-point amplitude and a disc one-point amplitude respectively. More details of this analogy are described in appendix E.

In the case of a finite interval, different spacetime metrics are classified up to diffeomorphisms (which are required to fix the endpoints) simply by the total proper time T , which is the integral of the lapse $T = \int N$. The path integral over matter fields on this fixed metric computes matrix elements $\langle\psi_2|e^{-iHT}|\psi_1\rangle$ of time evolution between initial and final states. But this leaves us with the gravitational path integral over all metrics, which here is simply the integral over the ‘modulus’ T . The natural measure for this integral is the obvious flat measure dT [37], but we also have to decide on a range of integration. A full exploration of the possibilities is in [27]. The only one which leads to a good pairing between physical states (i.e., which only depends on the co-invariant cosets $|\psi_{1,2}\rangle\rangle$ of $|\psi_{1,2}\rangle$) is to integrate over the whole real line, including positive and negative T . The upshot is that the path integral over all spacetimes between initial and final boundaries at finite time lands us on

the group averaging integral (4.5) (up to an arbitrary normalisation which can be absorbed into the measure):

$$\langle\langle \psi_2 | \psi_1 \rangle\rangle_{\text{GA}} = \int \frac{\mathcal{D}g \mathcal{D}q}{\text{Diffs}} e^{iS}. \quad (4.17)$$

Such a relation formally holds also in higher dimensions, and the path integral in fact provides a reasonable definition for the group average in gravity. We can interpret this path integral as an integral over the gauge group, which consists of diffeomorphisms of spacetime modulo diffs which vanish on a specific fixed t Cauchy surface where a wavefunction is defined (in this case, the initial or final Cauchy surface corresponding to the end of the interval). See further discussion of this path integral in [17]. Finally we note that in the BRST formalism, the integral over an unfixed modulus is associated with insertion of a b -ghost (see [38], for example), which is precisely what we require to get a nonzero pairing $\langle \uparrow | b | \uparrow \rangle$ between co-invariant states in the ghost sector.

For the last possibility we consider pairings between invariants, thought of as a path integral where spacetime is the whole real line (infinite in both directions). In this case, there is a single diffeomorphism class of metrics, so we have no moduli. However, after fixing to the gauge $N = 1$ where coordinate and proper time agree (for example), we are still left with a one-parameter family of residual diffeomorphisms: the rigid time translations. The infinite norm $\propto \delta(E = 0)$ of invariants using the inner product on the unconstrained Hilbert space \mathcal{H}_0 arises from the volume of this residual gauge group. In higher dimensions, the analogue is that we choose a local gauge condition which determines classical evolution completely (e.g., synchronous gauge or harmonic gauge), but we are still left with the residual diffeomorphisms corresponding to choice of $t = 0$ slice (and coordinates on that slice). Note that this residual gauge group is precisely the group described above (diffs modulo those which fix $t = 0$), which is also the space of possible Cauchy surfaces (including choice of spatial coordinates on the surface) in a given spacetime.

We see two possible ways to deal with the residual gauge group and get a finite result for the path integral. The first is to directly divide out by the infinite volume of the gauge group: this is what the spectral analysis inner product does, but we do not know how to generalise this beyond our simple 1D gravity models (it seems likely that the infinite volume of the unfixed gauge group would be field-dependent in higher dimensional models, if one could make sense of it). The more obvious alternative is to implement a further gauge condition which selects the $t = 0$ slice (and in higher dimensions, coordinates on that slice to fix momentum constraints).

For one possible class of gauge conditions, we can select a clock field $\chi(q)$ (depending only on target space coordinates q for now) and fix the $t = 0$ slice as a moment of time when χ vanishes, setting $\chi(q(t = 0)) = 0$. This means inserting $\delta(\chi(q(0)))$ in the path integral. But to make the result independent of gauge choice, we should also insert a Fadeev-Popov measure factor, which is the Jacobian determinant of the derivatives of the gauge-fixing functions with respect to gauge transformations. For our one-dimensional models, this ‘determinant’ is just the proper time derivative $\dot{\chi}$, so with this gauge choice we insert $\delta(\chi)\dot{\chi}$ at $t = 0$ into the path integral. Using the classical equations of motion, the measure factor becomes $\dot{\chi} = p_N$, the component of momentum normal to the surface $\chi = 0$. Quantum

mechanically, $\delta(\chi)$ and p_N do not commute so we must choose an ordering prescription to define the quantum operator corresponding to our insertion. The average (or Weyl) ordering $\frac{1}{2}(\delta(\chi)p_N + p_N\delta(\chi))$ is a reasonable prescription (which we will later justify in the BRST formalism). This operator is precisely our Klein-Gordon form κ_{KG} (4.8). We can see this most directly by noting that our insertion $\delta(\chi)\dot{\chi}$ can also be written as the time derivative of $\Theta(\chi)$, and comparing to our expression (4.15). This connection between the Klein-Gordon form and gauge fixing was noted in [32] and discussed more recently in [15]. In string theory, essentially the same calculation makes sense of sphere/disk amplitudes with two closed/open string vertex operators respectively [39] (see appendix E for more on the analogy with the string worldsheet).

The path integral interpretations of group averaging and Klein-Gordon inner products give us a nice interpretation of our compatibility condition $\eta\kappa\eta = \eta$ for inner product on invariants and co-invariants, shown in figure 2. The right hand side η represents a path integral over all spacetimes computing a group average inner product. As usual in the path integral formalism, Hilbert spaces arise when we cut a path integral, dividing it into a ‘ket’ piece computing one wavefunction at a given time and a ‘bra’ piece computing the other wavefunction, while the inner product sews these pieces back together. This gives us an interpretation of $\eta\kappa\eta$ on the left hand side, where each piece of the path integral gives a factor of η thought of as preparing an invariant wavefunction, and κ represents the sewing procedure.¹⁹ This ability to cut and sew back together embodies a notion of locality, and positivity of the inner product embodies unitarity (allowing a sensible probabilistic interpretation), and these are the reasons why Hilbert spaces are important. For QFT in a fixed background we can simply cut along any Cauchy surface Σ that we choose, and sewing back together simply requires identifying fields on Σ and integrating over them: this is a simple ‘ L^2 inner product’ of fields on Σ . But naively following this in gravity leads to a huge overcounting (responsible for the divergent nature of η^2), because each history will be counted once for every possible cutting surface. We have already seen the resolution: we must choose a gauge condition to select a unique slice Σ for every spacetime appearing in the path integral, leading to the insertion of κ as our gluing inner product.

However, there is one significant problem with using the Klein-Gordon form for such a gauge fixing (already noted in [32]), which is responsible for a disagreement between Klein-Gordon and group averaging inner products (so $\eta_{\text{GA}}\kappa_{\text{KG}}\eta_{\text{GA}} \neq \eta_{\text{GA}}$). The problem is that the Fadeev-Popov determinant $\dot{\chi}$ does not always give the correct measure factor: it is correct only up to a sign, and really we should take the absolute value. The upshot is that we expect compatibility between group averaging and Klein-Gordon inner products in cases where $\dot{\chi} > 0$ when crossing the gauge surface $\chi = 0$, and discrepancy by a sign when $\dot{\chi} < 0$ there. We could also have situations where $\chi = 0$ is not a complete gauge fixing, with a particular spacetime configuration crossing the gauge surface several times (or not at all). In such cases, κ_{KG} will count the number of crossings with signs (necessary for invariance under deformations of χ , which will create or destroy $\chi = 0$ crossings in cancelling pairs).

¹⁹We have given an interpretation in terms of invariant states created by group averaging some initial and final co-invariant wavefunctions. An alternative interpretation of the cutting in terms of co-invariants is that κ is a completeness relation, inverse to the matrix of inner products η on co-invariants.

Semi-classically, we can simply deal with this deficiency by hand: no problems will arise in perturbation theory around a classical solution for which $\chi = 0$ is a good gauge, and this is sufficient for most applications. But we do not expect that inserting the Jacobian with an absolute value $|\dot{\chi}|$ by hand (as suggested in [40], for example) will make sense in an exact quantum theory (since in the UV, we expect the typical path to become very jagged like Brownian motion, crossing the $\chi = 0$ slice infinitely many times).

Once we have this gauge-fixing interpretation there is nothing particularly special about Klein-Gordon inner products: we can choose χ to be any real function of q and p (or any Hermitian operator). An example is a zero-momentum gauge $\chi = p_0$ introduced in section 3.6.1 in analogy to the maximal volume gauge $\chi = \text{Tr } K$, for which [15] gives an extensive discussion of the gauge fixing and Fadeev-Popov determinant. Our choice will be guided by our desire that $\chi = 0$ gives a good gauge, preferably with all relevant trajectories in the path integral crossing the gauge slice in the correct direction from $\chi < 0$ to $\chi > 0$ so no wrong sign appears in the Fadeev-Popov measure. But we expect that no choice will satisfy this universally (gravity has an insoluble Gribov problem); the best we can do is make a case-by-case choice which works semi-classically for a particular problem.

4.5 Example: a one-dimensional target space

To make the above constructions more explicit and compare the resulting inner products, we can look at the simple example of one-dimensional target spaces. In such cases, our sigma-model Hamiltonian (2.4) can be written as

$$H = -\frac{1}{2}p^2 + V(q). \quad (4.18)$$

This is the familiar non-relativistic particle in a potential, apart from the ‘wrong sign’ kinetic term (from q being a ‘timelike’ coordinate with a one-dimensional Lorentzian metric $G_{00} = -1$). Take care to note that because of this, momenta and velocities are related with a sign, $\dot{q} = -p$.

Note that examples with higher-dimensional target space can also reduce to this form if the ‘timelike’ and ‘spacelike’ dependence factorise. A familiar example takes \mathcal{M} to be Minkowski spacetime with constant potential $\frac{1}{2}m^2$; this is the worldline theory of free massive QFT. Working in a basis of fixed spatial momentum k reduces this to a family of decoupled effectively one-dimensional Hamiltonians (4.18) with constant $V(q) = \frac{1}{2}\omega^2$ with $\omega^2 = |k|^2 + m^2$.

For connections to gravity, a possible interpretation to have in mind (borne out in mini-superspace examples) is to interpret e^q as a volume of space, so movement to the left/right corresponds to contraction/expansion of the universe.

First consider the simple case of a constant potential $V(q) = \frac{1}{2}\omega^2$ with $\omega > 0$ (which includes the Minkowski example as described above). Translation-invariance of H makes the states very simple to describe in momentum space. Co-invariant states are given by (smooth, say) momentum-space wavefunctions $\tilde{\psi}(p)$ modulo functions of the form $(p^2 - \omega^2)f(p)$, which are simply the wavefunctions that vanish at $p = \mp\omega$. So the coset of $\tilde{\psi}(p)$ is determined by its values $\tilde{\psi}(\mp\omega)$ at those points, and \mathcal{H}_{co} is two-dimensional. Invariant

states are linear combinations of $\tilde{\psi}_{\pm}(p) \propto \delta(p \pm \omega)$, or of $\psi_{\pm}(q) \propto e^{\mp i\omega q}$ in position space: positive- and negative-frequency solutions to the Wheeler-DeWitt equation $H\psi = 0$ (noting that positive frequency means negative p and vice-versa, due to the negative kinetic term).

Using the rigging map $\eta = 2\pi\delta(H)$ directly gives

$$\begin{aligned}\eta &= 2\pi\delta\left(\frac{1}{2}(p^2 - \omega^2)\right) = \frac{2\pi}{\omega}(\delta(p - \omega) + \delta(p + \omega)), \\ \langle q|\eta|q'\rangle &= \frac{2}{\omega}\cos(\omega(q - q')).\end{aligned}\tag{4.19}$$

The position space result in the second line is a Fourier transform of the momentum space expression, or can be obtained by the (conditionally convergent) group average integral of the free non-relativistic propagator $\langle q'|e^{-iHt}|q\rangle$ over $t \in \mathbb{R}$. With this inner product, an obvious orthonormal basis $\{|+\rangle, |-\rangle\}$ of co-invariants is given by the ‘positive frequency’ coset $|+\rangle$ of momentum-space wavefunctions with $\psi(p = -\omega) = \sqrt{\frac{\omega}{2\pi}}$ and $\psi(p = +\omega) = 0$, and the ‘negative frequency’ coset $|-\rangle$ with $\psi(p = -\omega) = 0$ and $\psi(p = +\omega) = \sqrt{\frac{\omega}{2\pi}}$. Under η , these cosets map to invariants $|\pm\rangle = 2\pi\delta(H)|\pm\rangle$ with (position space) wavefunctions $\psi_{\pm}(q) = \frac{1}{\sqrt{\omega}}e^{\mp i\omega q}$, which (by definition) is an orthonormal basis of the dual inner product on invariants. This agrees with the spectral analysis inner product as expected: $\psi_{\pm}(q)$ are the $E = 0$ value of $\psi_{E,\pm}(q) = \frac{1}{\sqrt{|p|}}e^{ipq}$, where $p = \mp\sqrt{\omega^2 - 2E}$. These are eigenstates of H with eigenvalues E (for $E < \frac{1}{2}\omega^2$), normalised by $\langle \psi_{E',\epsilon'}|\psi_{E,\epsilon}\rangle = 2\pi\delta(E - E')\delta_{\epsilon\epsilon'}$ (with $\epsilon, \epsilon' = \pm$ specifying the sign of p). Note that this differs from the usual momentum basis by $\sqrt{|p|}$ in the denominator, because we normalise by delta-functions in E (not p).

The Klein-Gordon inner product on invariants for one-dimensional target spaces is

$$\llbracket \psi_2|\psi_1 \rrbracket_{\text{KG}} = -\frac{i}{2}(\psi_2^*(q)\psi_1'(q) - \psi_2'^*(q)\psi_1(q)),\tag{4.20}$$

noting that the negative kinetic term means that the normal derivative defining Π in (4.7) is $\partial_N = -\partial_q$. This is the Wronskian of a pair of solutions ψ_1, ψ_2^* to $H\psi = 0$, so is independent of q (for any potential V). For our simple constant V example, we find $\llbracket \pm|\pm \rrbracket = \pm 1$ (and $\llbracket \pm|\mp \rrbracket = 0$), so negative frequency solutions have negative Klein-Gordon norm. The corresponding form on co-invariants is

$$\begin{aligned}\eta_{\text{KG}} &= 2\pi\frac{\delta(p + \omega) - \delta(p - \omega)}{\omega} = -2\pi\text{sgn}(p)\delta(H), \\ \langle q|\eta_{\text{KG}}|q'\rangle &= \frac{2\sin(\omega(q - q'))}{i\omega}.\end{aligned}\tag{4.21}$$

We have written η_{KG} as $-\text{sgn}(p)$ times the group averaging rigging map to emphasise that they differ only by a sign: this is precisely the sign in the measure we commented on in the previous subsection. The position space result may be familiar (up to field normalisation) as the commutator $[\Phi(q), \Phi(q')]$ of a harmonic oscillator ‘position’ variable Φ at different values of target space ‘time’ q (since this is the target space one-dimensional ‘QFT’ corresponding to $H = \frac{1}{2}(\omega^2 - p^2)$, see sections 2.3 and 4.6).

We can generalise this analysis to a potential which is constant only asymptotically in the distant past and future, $V(q) \rightarrow \frac{1}{2}\omega^2$ for $|q| \rightarrow \infty$ [41]. We will summarise the upshot in

terms of invariants. There is still a two-dimensional space of invariants, but now there are two different natural bases of solutions to $H|\psi\rangle = 0$. These are in states $|\text{in}, \pm\rangle$ defined as a solution coming in from $q \rightarrow \pm\infty$ in the far past on the worldline, and out states $|\text{out}, \pm\rangle$ similarly going to $q \rightarrow \pm\infty$ in the future: see section 3.4. The asymptotics of the in state wavefunctions are as follows:

$$\begin{aligned} \psi_{\text{in},+}(q) &\sim \begin{cases} \frac{1}{\sqrt{\omega}} (e^{i\omega q} + S_{++} e^{-i\omega q}) & q \rightarrow +\infty \\ \frac{1}{\sqrt{\omega}} S_{-+} e^{i\omega q} & q \rightarrow -\infty \end{cases} \\ \psi_{\text{in},-}(x) &\sim \begin{cases} \frac{1}{\sqrt{\omega}} S_{+-} e^{-i\omega q} & q \rightarrow +\infty \\ \frac{1}{\sqrt{\omega}} (e^{-i\omega q} + S_{--} e^{i\omega q}) & q \rightarrow -\infty \end{cases} \end{aligned} \quad (4.22)$$

The out-states are given by complex conjugation (time-reversal) of out-states, $\psi_{\text{out},\pm} = \psi_{\text{in},\pm}^*$. Either of these bases are orthonormal for the group average (spectral analysis) inner product, and the two bases are related by the 2×2 unitary S-matrix $S_{\pm\pm}$ of our 1D quantum mechanics (which is also symmetric, $S_{+-} = S_{-+}$, due to time-reversal symmetry):

$$\langle\langle \text{out}, \epsilon' | \text{out}, \epsilon \rangle\rangle_{\text{GA}} = \delta_{\epsilon\epsilon'}, \quad \langle\langle \text{in}, \epsilon' | \text{in}, \epsilon \rangle\rangle_{\text{GA}} = \delta_{\epsilon\epsilon'}, \quad \langle\langle \text{out}, \epsilon' | \text{in}, \epsilon \rangle\rangle_{\text{GA}} = S_{\epsilon'\epsilon}. \quad (4.23)$$

The Klein-Gordon inner product is simplest to express in a mixture of the two bases, as overlaps between in-kets and out-bras:

$$\langle\langle \text{out}, \pm | \text{in}, \pm \rangle\rangle_{\text{KG}} = 0, \quad \langle\langle \text{out}, \pm | \text{in}, \mp \rangle\rangle_{\text{KG}} = \pm \langle\langle \text{out}, \pm | \text{in}, \mp \rangle\rangle_{\text{GA}}. \quad (4.24)$$

We can interpret this as saying that the Klein-Gordon form gives a vanishing contribution for trajectories which ‘reflect’, coming from $q = +\infty$ in the distant past and leaving to $q \rightarrow +\infty$ in the future or the same for $q \rightarrow -\infty$. Trajectories which start at $q \rightarrow -\infty$ and ‘transmit’ to $q \rightarrow +\infty$ have positive contribution (equal to the result from group averaging), and those in the opposite direction contribute negatively. One way to express this is a relation (equivalent to that given in [41])

$$\langle\langle \psi' | \psi \rangle\rangle_{\text{KG}} = \langle\langle \psi' | \frac{1}{2}(\epsilon_{\text{in}} + \epsilon_{\text{out}}) | \psi \rangle\rangle_{\text{SA}}, \quad (4.25)$$

where $\epsilon_{\text{in},\text{out}}$ are operators which give $\text{sgn}(p)$ in the asymptotic future or past (with in and out states as eigenstates respectively). These don’t commute (because an incoming wave from one direction scatters into a superposition of both out states).

These results could have been anticipated from the gauge-fixing interpretation given for the Klein-Gordon form in section 4.4. The two inner products agree when a constant q gauge-fixing surface is crossed once in a positive direction, they disagree with a sign when it is crossed once in a negative direction so the Jacobian is negative, and the Klein-Gordon form gives zero when a gauge-fixing surface is not crossed at all (or crossed an equal number of times in each direction). This interpretation is clear only when we compute overlaps between in- and out-states as above, since the contributing trajectories all begin and end in a definite direction of target space: the Klein-Gordon overlap between two in-states, for example, is more murky (until we re-express one in a basis of out states).

Another interesting case is when $V = \frac{1}{2}\omega^2$ for $q \rightarrow -\infty$ as before, but $V < 0$ for $q \rightarrow \infty$. This might correspond to ‘bang-crunch’ cosmologies that expand and recollapse. In this case, the allowed space of invariant states (defined by the image of η starting from smooth compactly supported wavefunctions, or by requiring that the form κ is well-defined) is quite different depending on the choice of η or κ . Group averaging (or spectral analysis) now gives a single state, corresponding to the solution of the Wheeler-DeWitt equation that decays at $q \rightarrow \infty$. This leads to physics in accord with semi-classical expectations [13], exponentially suppressing classically-forbidden configurations with large q . However, in the context of the Klein-Gordon norm it is natural to allow any initial data and hence any solution to $H|\psi\rangle = 0$ (the usual requirement is compactly supported initial data, which is vacuous in 1D), leading to a two-dimensional space of states. Alternatively, it has been suggested to keep only the solution which is purely positive frequency at $q \rightarrow -\infty$ [12, 42]. In either case, we include solutions which grow at $q \rightarrow \infty$ (in the associated target-space QFT described in the next section, this would correspond to uncontrolled tachyonic particle production).²⁰ This would appear (depending on the interpretation of the wavefunction) to favour large q , where there are no classical solutions, which seems challenging for a sensible semi-classical limit. We view this as further support for group averaging.

4.6 The QFT inner product

The motivation of DeWitt [7] and [12] for proposing the Klein-Gordon inner product was by analogy to quantum field theory. Specifically, as explained in section 2.2, the Hamiltonian we are considering arises as the equation of motion of a free field $\Phi(q)$ (which we take to be real here) on the manifold \mathcal{M} , with Lagrangian $\frac{1}{2}(\nabla\Phi)^2 + V\Phi^2$. Here, we explain the connection between this QFT and a quantum mechanical ‘worldline theory’ with the constraint $H = 0$ in our language of invariants and co-invariants. We will not use this in the remainder of the paper, so the reader can safely skip this part.

This connection depends on first choosing a ‘vacuum state’ $|\Omega\rangle_{QFT}$ of the QFT Hilbert space. In free QFT, it is natural to take $|\Omega\rangle_{QFT}$ to be a Gaussian state, meaning that the one-point function of Φ vanishes and the two-point (Wightman) function

$$W_\Omega(q_2, q_1) := \langle \Omega | \Phi(q_2) \Phi(q_1) | \Omega \rangle_{QFT} \quad (4.27)$$

determines all higher-point functions by a sum over Wick contractions. We can then define a ‘single particle’ Hilbert space $\mathcal{H}_{SP} \subset \mathcal{H}_{QFT}$ by acting on $|\Omega\rangle_{QFT}$ with an operator linear in the field Φ ,

$$|f\rangle_{QFT} := \int dq f(q) \Phi(q) |\Omega\rangle_{QFT} \in \mathcal{H}_{SP}. \quad (4.28)$$

²⁰The target space interpretation (where q is interpreted as time in target space) is a ‘free QFT’ in one spacetime dimension, with a scalar field Φ and time-dependent Hamiltonian

$$H_{\text{target}}(q) = \frac{1}{2}\Pi(q)^2 + V(q)\Phi(q)^2, \quad (4.26)$$

where Π is the momentum conjugate to Φ . This is a Harmonic oscillator with time-dependent frequency $V = \frac{1}{2}\omega^2$.

The n particle Hilbert space created by acting with a product of n such operators (minus Wick contractions) is then the symmetric product $\text{Sym}^n \mathcal{H}_{SP}$, and the full QFT Hilbert space is the Fock space given by the direct sum of all these sectors, $\mathcal{H}_{QFT} \simeq \bigoplus_{n=0}^{\infty} \text{Sym}^n \mathcal{H}_{SP}$. In this construction there is a great deal of freedom in the choice of vacuum state $|\Omega\rangle$, and hence in the subspace \mathcal{H}_{SP} of single particle states. Generically (in the absence of target space time-translation symmetry for example) there is no preferred vacuum, due to the familiar phenomenon of particle production. Different vacua are related by Bogoliubov transformations, which are the QFT interpretation of the S-matrix of the worldline theory (for example, the coefficients in (4.22) above).

Once the choice of $|\Omega\rangle_{QFT}$ has been made, we can relate the single particle Hilbert space \mathcal{H}_{SP} to the constrained state spaces V_{co} , V_{inv} . First, note that the single-particle states constructed in (4.28) are naturally associated with co-invariants $|f\rangle\rangle$, since the field obeys the equation of motion $H\Phi(q) = 0$. Taking $f = Hg \in \text{Im } H$, integrating by parts and using the equation of motion tells us that $\int f\Phi = \int gH\Phi = 0$, so all states $|f\rangle\rangle$ in the same co-invariant coset lead to the same QFT state. Second, note that we can identify a single particle state $|\Psi\rangle_{QFT}$ with an invariant state $|\psi\rangle$ by taking the wavefunctions to be matrix elements of the field operator, $\psi(q) = \langle\Omega|\Phi(q)|\Psi\rangle$. This gives a nonzero wavefunction for any single particle state of \mathcal{H}_{QFT} , and zero for other multi-particle sectors. We can summarise this by the following linear maps:

$$V_{\text{co}} \rightarrow \mathcal{H}_{SP} : |f\rangle\rangle \mapsto |f\rangle_{QFT} := \int_{\mathcal{M}} dq f(q)\Phi(q)|\Omega\rangle_{QFT} \quad (\text{surjective}), \quad (4.29)$$

$$\mathcal{H}_{SP} \rightarrow V_{\text{inv}} : |\Psi\rangle_{QFT} \mapsto |\psi\rangle, \quad \psi(q) = \langle\Omega|\Phi(q)|\Psi\rangle_{QFT} \quad (\text{injective}). \quad (4.30)$$

These maps are ‘dual’, meaning that they relate the canonical pairing between V_{co} and V_{inv} and the QFT inner product:

$$\langle f|\Psi\rangle_{QFT} = \int dq f^*(q)\langle\Omega|\Phi(q)|\Psi\rangle = \langle\langle f|\psi\rangle\rangle. \quad (4.31)$$

From this, the QFT inner product on \mathcal{H}_{SP} induces an inner product on V_{co} by pullback under the above linear map,

$$\begin{aligned} \langle\langle f_2|f_1\rangle\rangle_{\Omega} &= \langle f_2|f_1\rangle_{QFT} = \int dq_1 dq_2 f_2^*(q_2) f_1(q_1) W_{\Omega}(q_2, q_1) \\ &\implies \eta_{\Omega} = W_{\Omega}. \end{aligned} \quad (4.32)$$

This identifies the map η_{Ω} from co-invariants to invariants as the Wightman function W_{Ω} . This inner product is positive semi-definite, but unlike the examples of η discussed above it is degenerate. This means that the image of η_{Ω} (the space of invariant wavefunctions $|\psi\rangle$ that arise from a state $|\Psi\rangle_{QFT}$ in the QFT) does not cover all invariants. Instead, it will cover only half of V_{inv} . When $|\Omega\rangle_{QFT}$ is the vacuum state in a static target space these are the positive frequency wavefunctions. The resulting co-invariant Hilbert space contains additional null states besides $\text{Im } H$.

For the dual Hilbert space of invariants, for the inner product we can in fact use the Klein-Gordon form (4.8). Instead of modifying the inner product, we simply restrict the space of allowed states to a ‘positive frequency’ subspace $V_+ \subset V_{\text{inv}}$ on which κ_{KG} is

positive-definite. This selects ‘half the states’ since we must have $V_{\text{inv}} = V_+ \oplus V_-$ where the negative frequency space $V_- = V_+^*$ consists of complex conjugates of positive frequency solutions. The choice of V_+ is an alternative characterisation of the state $|\Omega\rangle_{QFT}$.

5 Operators on the physical Hilbert space

In quantum mechanics, Hermitian operators play two (closely related) roles: they are observables, and they generate unitary flows. The most prominent example of the latter is the Hamiltonian, which generates time translation. Gravity is particularly unusual in this respect since time translation is locally a gauge symmetry (though we can still consider physical time translation operators in the presence of timelike boundaries such as the conformal boundary in asymptotically AdS spacetimes). Nonetheless, in any gauge theory we must consider the constraints when we discuss operators and their flows on the physical Hilbert space. There are two complementary considerations, both classically and quantumly. First, we are constrained because we want the flow generated by a physical operator to map physical states (satisfying constraints) to physical states. On the other hand, we have additional freedom because we can always add a gauge transformation to a flow, leading to a different but physically equivalent point on a gauge orbit.

Commonly, the criterion given for physical operators is that they are gauge-invariant, commuting (or Poisson commuting in the classical theory) with the constraints. This is convenient (and perhaps sufficient) for observables, and it certainly suffices to satisfy the above constraint of mapping physical states to physical states (whether using the invariant or co-invariant Hilbert space). However, it is unnecessarily and overly restrictive, particularly for operators generating flows like Hamiltonians, since it does not allow us the freedom of a gauge choice. A simple example is $U(1)$ gauge theory, where the usual Hamiltonian contains terms of the form $E^i \partial_i A_0$ which do not commute with the primary constraint Π^0 (the momentum conjugate to A_0), along with possible additional gauge-fixing terms: see appendix F for details of this example. This is even more true in gravity, where the non-abelian algebra of Hamiltonian constraints makes it impossible to write an operator which generates time-translations and also commutes with constraints.

Given this motivation, in this section we review the more general conditions on physical operators, and comment on the action of the resulting physical operators on the Hilbert spaces of invariants and co-invariants discussed above. We do this first classically in the language of Dirac, and then explain how this relates to the BRST formalism (which provides a powerful and convenient way to ensure gauge invariance in the quantum theory). Of particular importance for us will be the trivial operators which act non-trivially on the unconstrained Hilbert space \mathcal{H}_0 but are physically equivalent to the zero operator. These can be interpreted as infinitesimal quantum gauge transformations. In the BRST formalism we can exponentiate these trivial (BRST-exact) operators to finite gauge transformations, which we will use in the next section to construct inner products in a more systematic manner.

5.1 Dirac observables

We first review observables in theories with constraints in Dirac's formalism [6], starting with the classical theory in which observables are functions of phase space.

First, we have a constraint for a function $O(q,p)$ of phase space to be a physical observable: it should be constant on gauge orbits of physical states. This means that Poisson brackets of O with the constraints (i.e., the gauge variation of O) should vanish *weakly*, meaning they are zero on the constraint surface. In our case with a single constraint H , this means that $\{H, O\} = H\tilde{O}$ for some function \tilde{O} (so the time derivative of O vanishes if the energy is zero). This also means that the flow generated by O preserves the constraint surface, and that gauge-equivalence is retained by the flow (i.e., two points starting on the same gauge orbit end on the same gauge orbit).

In addition to this constraint, we have some freedom. Two different functions of phase space correspond to the same physical observable if they are weakly equal (equal on the constraint surface), meaning we have the equivalence $O \sim O + H\xi$ for any ξ . Note that if O is a good observable then so is $O + H\xi$, since $\{H, H\xi\} = H\{H, \xi\}$ is weakly zero.²¹

Quantum observables are similar, though we need to take operator ordering into account. Depending on whether we take our observables to act on invariant or co-invariant states, different choices of ordering are needed. The conditions for an operator O on \mathcal{H} to give rise to a well-defined operator on the constrained Hilbert space, and the equivalences between such O , are the following:

$$\begin{aligned} O \text{ operator on } \mathcal{H}_{\text{inv}} : & \quad [H, O] = i\tilde{O}H, & \quad O \sim O + \xi H, \\ O \text{ operator on } \mathcal{H}_{\text{co}} : & \quad [H, O] = iH\tilde{O}, & \quad O \sim O + H\xi. \end{aligned} \tag{5.1}$$

The condition in the first line ensures that $HO|\psi\rangle = 0$ whenever $H|\psi\rangle = 0$, and the equivalence arises because ξH annihilates all invariants. Similarly, the condition in the second line is required so that O maps null states to null states, and the equivalence comes from identifying states in the same co-invariant coset. Note that these observables are in one-to-one correspondence: if O_{inv} gives a well-defined operator on invariants (with $[H, O_{\text{inv}}] = i\tilde{O}H$ as above), then $O_{\text{co}} = O_{\text{inv}} + i\tilde{O}$ is well-defined on co-invariants.

If we have several constraints, we can continue to define observables as equivalence classes of operators which weakly commute with constraints modulo weakly zero operators. For the quantum theory the ordering in these definitions is appropriate to invariance or co-invariance as in (5.1), but may use different prescriptions for different constraints. For example, in the non-minimal formalism introduced at the end of section 3.3, invariant states with respect to H are treated as co-invariant with respect to Π and vice-versa.

In such cases with several constraints and states with a mixture of invariance and co-invariance, this definition of observables has a significant drawback: it does not give us a well-defined algebra. For example, the product of O with a weakly zero operator may not be weakly zero, even if O weakly commutes with constraints. Taking the specific example of states invariant with respect to H and co-invariant with respect to Π , both H and ΠA (for any matter operator A) are weakly zero, but their product $H\Pi A$ may not

²¹More generally, this is true for any set of first class constraints.

be if $[H, A]$ is non-zero. In particular, we cannot guarantee the unitary flow constructed by exponentiating these operators will be gauge-invariant. This is resolved in the BRST formalism.

Before moving onto BRST observables, we make one comment on the narrower class of observables which exactly commute with constraints, for which none of these complications from operator ordering arise.²² One class of examples is constructed by averaging, such as $O = \int dt e^{+iHt} O_0 e^{-iHt}$ for any O_0 (such that the integral converges), which in the path integral might be written as the integral of a local quantity over all of spacetime [43, 44]. Such operators can act both on invariants and co-invariants. It is reasonable to demand that these two actions should be physically equivalent, meaning that O commutes with the rigging map η . Imposing this requirement essentially uniquely selects η to be the group average [33], lending further support to this proposal.

5.2 Observables in BRST

The description of physical observables in section 5.1 is reminiscent of cohomology: a requirement to weakly commute with constraints is reminiscent of closedness, while the equivalence up to weakly zero operators is reminiscent of exactness. This is realised precisely in the BRST formalism, where the physical operators are given by the operator cohomology (at ghost number zero).

Classically, consider functions \mathbf{O} on the BRST extended phase space (depending on Grassmann ghost variables b, c as well as q, p) with $N_g = 0$. The ghost-free part of \mathbf{O} (obtained by setting ghosts to zero) is an ordinary function O of bosonic variables q, p which will correspond to the ‘Dirac’ observable discussed above. The Dirac condition that Poisson brackets of O with the constraints are weakly zero is replaced by the condition that the Poisson bracket of \mathbf{O} with the BRST charge is *exactly* zero, $\{Q, \mathbf{O}\} = 0$ (or \mathbf{O} is BRST-closed). The equivalence between weakly equal operators is replaced with triviality of BRST-exact operators, defined as the Poisson bracket $\{Q, \Psi\}$ for any function Ψ with $N_g = -1$. Hence the physical operators are elements of the BRST cohomology of closed \mathbf{O} modulo exact \mathbf{O} at $N_g = 0$.

To see how this works, let’s explicitly work through our case of a single constraint with $Q = cH$. We can write the most general $N_g = 0$ function as $\mathbf{O} = O - ibc\tilde{O}$ with O, \tilde{O} independent of ghosts, and we have $\{Q, \mathbf{O}\} = c(\{H, O\} - \tilde{O}H)$.²³ From this we see that O weakly commutes with the constraint if and only if it has a BRST-closed extension \mathbf{O} . For the BRST-exact operators we consider the Poisson bracket of Q with the most general $N_g = -1$ operator $\Psi = b\xi$, giving $\{Q, \Psi\} = -i(H\xi - ibc\{H, \xi\})$; this means we can identify $O = H\xi$ (with $\tilde{O} = \{H, \xi\}$) as a trivial observable.

²²One may have to take more care when there are field-dependent structure constants as occurs in gravity, since then an operator might commute with the constraints but not with the BRST charge.

²³The Poisson bracket on the extended phase space is given by

$$\{f, g\} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} - i \left(\frac{\partial^R f}{\partial b} \frac{\partial^L g}{\partial c} + \frac{\partial^R f}{\partial c} \frac{\partial^L g}{\partial b} \right), \quad (5.2)$$

where the L, R in the ghost terms indicate left and right derivatives. This gives the usual correspondence between quantum (anti-)commutators and Poisson brackets, $[\cdot, \cdot]_{\pm} \sim i\{\cdot, \cdot\}$.

For more complicated cases with many constraints (particularly for non-abelian algebras and field-dependent structure constants), the Dirac condition of weakly commuting with constraints is not quite enough to get a physical observable. The requirement that O can be completed to a BRST exact \mathbf{O} by the addition of ghost terms gives the Dirac condition, but also imposes further conditions at higher order in ghosts.

Quantum mechanically, we simply replace Poisson brackets in the above discussion with (anti-)commutators. The BRST formalism then automatically takes care of the different ordering prescriptions discussed above. A BRST observable $\mathbf{O} = O - ibc\tilde{O}$ is closed ($[Q, \mathbf{O}] = 0$) if $[H, O] = iH\tilde{O}$ as appropriate for co-invariants in (5.1); and indeed co-invariants are accompanied by ghost states annihilated by c , so \mathbf{O} acts on these states simply as the ghost-free part O . Similarly, a BRST-exact operator $[Q, b\xi]_+ = H\xi + bc[\xi, H]$ generates the equivalences appropriate for co-invariants. On the other hand, on invariant states (with ghosts annihilated by b) the same operator \mathbf{O} will instead act as $O_{\text{inv}} = O - i\tilde{O}$, giving the correspondence between operators acting on invariants and co-invariants discussed after (5.1). The BRST observable \mathbf{O} packages both possibilities (and all possible mixtures if there are more constraints) into a single object.

Furthermore, in this formalism the physical observables in fact form an algebra: the product of two BRST-closed operators is closed, and the product of a BRST-closed (physical) operator with a BRST exact (trivial) operator is BRST exact (trivial). This follows simply from the super-Jacobi identity: for example, $[Q, \mathbf{O}_1\mathbf{O}_2] = [Q, \mathbf{O}_1]\mathbf{O}_2 + \mathbf{O}_1[Q, \mathbf{O}_2] = 0$ if $\mathbf{O}_{1,2}$ are closed. An important consequence is that this allows us to discuss the unitary flow generated by a Hermitian operator O : the exponential $e^{is\mathbf{O}}$ is a physical (BRST closed) operator, with the extra ghost terms in the operator providing the Hamiltonian version of the Fadeev-Popov determinants familiar from the path integral. An important special case which we will use in the next section is a finite quantum gauge transformations, meaning the exponential of a BRST exact operator $\exp[Q, \Psi]_+$, which is BRST-equivalent to the identity.

6 Inner products from the BFV-BRST formalism

In section 4 we discussed general properties of inner products on \mathcal{H}_{inv} and \mathcal{H}_{co} , along with some examples. From that discussion alone there is a great deal of freedom in defining an inner product, and the choice seems rather arbitrary. In this section we describe a way to obtain inner products in a more systematic way from the BFV-BRST formalism [21, 22], which removes much of that freedom.

Useful references for general tools on the BFV-BRST formalism include [23] and the book [24]. The main ideas for the application to inner group averaging follow [18–20].

6.1 Sketch of the main ideas

For this, we will use the non-minimal BFV version of the BRST formalism introduced around equation (3.7), where we include the lapse N as a dynamical degree of freedom compensated by the constraint $\Pi = 0$, requiring that the momentum conjugate to N vanishes. One way to motivate this choice is that now we can choose the physical states to

reside at ghost number $N_g = 0$, which is necessary to get a non-zero inner product (since N_g is anti-Hermitian). But this doesn't work immediately; let's see what happens when we naïvely try to take the inner product of two Wheeler-DeWitt (invariant) states in the BFV Hilbert space (3.7):

$$|\psi\rangle \otimes |\psi_N\rangle \otimes |\uparrow\downarrow\rangle \in \mathcal{H}_0 \otimes \mathcal{H}_N \otimes \mathcal{H}_{\text{ghost}} \text{ with } H|\psi\rangle = 0 \longrightarrow \underbrace{\langle\psi'|\psi\rangle}_{=\infty} \langle\psi'_N|\psi_N\rangle \underbrace{\langle\uparrow\downarrow|\uparrow\downarrow\rangle}_{=0}. \quad (6.1)$$

We find an indefinite form: the first factor $\langle\psi'|\psi\rangle$ from the unconstrained Hilbert space gives infinity since invariant wavefunctions are non-normalisable, as remarked below equation (3.1). But this is multiplied by $\langle\uparrow\downarrow|\uparrow\downarrow\rangle = 0$ from the ghost sector.²⁴ The idea is to resolve this degeneracy by inserting a finite 'gauge transformation' $e^{[Q,\Psi]_+}$, the exponential of a BRST exact (and hence physically trivial) operator, in such a way that we get a finite result. We then hope to interpret this finite result as a physical inner product between invariants. Similar remarks apply to co-invariant states (with the infinity coming from the norm of the lapse wavefunction $|\Pi = 0\rangle$ annihilated by Π). We will see precisely how this works in examples below.

This construction has several important properties associated with BRST-invariance. The first is that it depends only on the BRST cohomology class of the states, since acting on a BRST exact state with $e^{[Q,\Psi]_+}$ gives another BRST exact state. This means that the result for invariant states should be independent of the lapse-dependence of wavefunctions (except for the integral over N), and for co-invariant states will depend only on the co-invariant coset as desired. The second property is that infinitesimal changes in Ψ do not affect the matrix elements of $e^{[Q,\Psi]_+}$ between BRST-closed states (the Fradkin-Vilkovisky theorem [21]). At least naïvely, this means that the result should be independent of the choice of Ψ . This means that there is not in fact much freedom in choosing an inner product; there is an essentially unique choice.²⁵

However, we will also show below how to obtain both group-average and Klein-Gordon inner products in this way, and these are different! More precisely (since they are inner products on different Hilbert spaces \mathcal{H}_{co} and \mathcal{H}_{inv}), they are not compatible by the $\eta\kappa\eta = \eta$ generalised inverse criterion. The resolution is a loophole to the Fradkin-Vilkovisky theorem that allows for discrete jumps as we vary Ψ , since the theorem does not tell us about finite changes in Ψ if the amplitude passes through an indeterminate form (most notably the original $\Psi = 0$ in (6.1)).

²⁴The lapse sector inner product $\langle\psi'_N|\psi_N\rangle$ could be finite, zero, or infinity depending on what representatives of the BRST cohomology class we choose, but in any case the result is indefinite.

²⁵This is true once we have chosen a BRST charge Q . For our simple one-dimensional gravity model this still leaves some freedom, for example by rescaling the Hamiltonian H by a function of fields to get a new classically-equivalent constraint \tilde{H} , and defining a different BRST operator $\tilde{Q} = c\tilde{H} + \dots$: a specific example is discussed in section 6.4.1. We regard this choice of Q as part of the ambiguity inherent in quantising a classical theory. Nonetheless, in richer models with non-abelian constraint algebras and spatial dependence we expect the requirement of a local nilpotent Hermitian BRST charge to be much more restrictive, so there will be less remaining ambiguity (in perturbation theory, hopefully only the usual ambiguity of renormalisation of parameters, including generating higher-derivative terms).

There is a more physical perspective on these algebraic manipulations related to gauge fixing. The original infinity we encountered in the inner product of invariants can be thought of as the infinite volume of the gauge group, as remarked in section 4.4. To get a finite result we need to fix this overcounting of equivalent configurations, so that we integrate over every gauge-orbit exactly once. In the construction we are discussing here, the bosonic part of the generator $[Q, \Psi]_+$ typically has precisely this effect, with a gauge condition appearing after integrating out Π and/or N (though with this BRST approach we never need to deal directly with the resulting second-class constraints, Dirac brackets and so forth). For this reason, Ψ is usually called the *gauge-fixing fermion*. The ghost terms in $[Q, \Psi]_+$ have the job of fixing an invariant measure on the resulting gauge-fixed slice, by contributing the Fadeev-Popov determinant Δ_{FP} .

However, as we already saw in section 4.4 for the example of the Klein-Gordon form, Δ_{FP} is not quite the correct measure, since it is not automatically positive and we should take its absolute value. Furthermore, we cannot simply fix this by choosing a sign in the final result, since Δ_{FP} can vanish and its sign can change as we move between different gauge orbits or change the gauge condition. This indicates that our choice of Ψ corresponds to a gauge condition which is not globally valid: a Gribov problem [45]. A given gauge-orbit may not have any representative satisfying the gauge condition, or it may have several (perhaps with different signs for Δ_{FP}). These phenomena give an intuitive explanation for the possibility of jumps as we vary the gauge-fixing fermion Ψ . It also suggests a criterion for selecting a unique physical inner product: it should arise from some choice of Ψ , give a positive measure to every gauge-orbit (leading to a positive semi-definite inner product), and count each orbit exactly once. These criteria are satisfied by the group-averaging proposal for our one-dimensional gravity models, which is perhaps the strongest evidence that this is the physically correct inner product. It is plausible that this remains true in more complicated theories, though there are significant technical obstacles to making this completely precise in realistic theories of gravity, some of which are discussed in section 8.

6.2 The non-minimal formalism and gauge-fixing

Before introducing the full BRST machinery, we first give a brief explanation of the reasons for using the non-minimal BFV formalism, and the connection to gauge-fixing. This will be helpful for interpreting the results below.

We first compare the classical variational principles coming from first-order actions, using the example of the Hamiltonian $\mathbf{H} = NH$ generating proper time evolution as given in (2.2) (obtained by a naïve canonical recipe from the action (2.1)), with and without including a momentum Π conjugate to N :

$$S_{\text{minimal}}[q, p, N] = \int dt [pq - NH(q, p)], \quad (6.2)$$

$$S_{\text{non-minimal}}[q, p, N, \Pi] = \int dt [pq + \Pi\dot{N} - NH(q, p)]. \quad (6.3)$$

Varying the ‘minimal’ action (in which Π is absent, and N acts only as a Lagrange multiplier to enforce $H = 0$) gives equations of motion which do not have a unique solution: N is completely undetermined. This is nothing but the gauge symmetry under diffeomorphisms

of time. But for the non-minimal action, the additional equation of motion from varying Π means that initial data for q, p, N, Π determine a unique time evolution. The extra variable Π acts as a Lagrange multiplier fixing a gauge condition, in this case the ‘synchronous’ gauge $\dot{N} = 0$.

This simple observation raises the question: can we modify the action to describe evolution in other gauges? The discussion in section (5.1) makes it clear that we can achieve this by adding terms to the Hamiltonian which are weakly zero, i.e., which vanish when constraints are satisfied. Specifically, by adding a term $F(q, p, N)\Pi$ linear in Π to get the full Hamiltonian $\mathbf{H} = NH + \Pi F$, we get a Hamiltonian that continues to generate time evolution but now in a much more general gauge (with $\dot{N} = F$). We can also add higher order terms in Π (e.g., a multiple of Π^2): this will not change the classical dynamics when on the constraint surface $\Pi = H = 0$, but can be useful in the quantum theory where this gives us the Gaussian-averaged gauges familiar from gauge theory.

Gravity is special in this context, because time evolution (except at spatial boundaries) is a gauge transformation: the Hamiltonian $\mathbf{H} = NH$ itself is weakly zero, so these flows simply move us between different points on the same gauge-orbit. In non-gravitational theories (or for gravity with asymptotic boundaries where time evolution is physical), the Hamiltonian will be a non-trivial Dirac observable as discussed in section (5.1), i.e. a function which weakly commutes with the constraints, ambiguous up to addition of a weakly zero function. The example of Maxwell theory is discussed in appendix F.

In gravity (or any other theory) we can consider gauge transformations generated by any weakly zero function $G(q, p, N, \Pi)$, which need not include a term NH allowing us to interpret the flow in terms of time translation. For example, for any $\chi(q, p)$ we can consider the flow generated simply by $G = \Pi\chi(q, p)$. On the constraint surface (with initial condition satisfying $\Pi = 0$), the classical evolution leaves q, p unchanged and relates the change in the lapse N to $\chi(q, p)$: in particular, classical solutions with equal initial and final values of lapse N exist only if $\chi = 0$. So under these conditions on initial and final N states, the transformation generated by $\Pi\chi(q, p)$ can be interpreted as imposing the gauge condition $\chi = 0$ at a single moment of time. This motivates using such a flow to impose gauge fixing on invariant states as discussed in section 4.4.

When we pass to the quantum theory, guaranteeing gauge-invariance is more subtle due to issues of operator ordering (or the measure in the path integral approach). In section 5, this showed up by the failure of Dirac observables to form an algebra. The BRST formalism resolves this: in place of a weakly-zero generator $G(q, p, N, \Pi)$ as in the examples discussed above we use a BRST exact operator $[Q, \Psi]_+$ for some gauge-fixing fermion Ψ . We get back the weakly zero function G from the ghost-free part of $[Q, \Psi]_+$ (in the classical limit when ordering becomes unimportant).

6.3 Non-minimal BRST

With these preliminaries, we are now ready to study the quantum theory in the non-minimal BRST formalism. First, we briefly recall the key features already introduced earlier. For the bosonic part of the Hilbert space, we have wavefunctions depending on the original unconstrained variables q as well as the real-valued lapse N . Associated with the

two constraints H and Π are two sets of ghosts c and \hat{c} and their conjugate momenta b and \hat{b} , with non-trivial anti-commutation relations $[c, b]_+ = [\hat{c}, \hat{b}]_+ = 1$ represented on a four-dimensional ghost Hilbert space with basis $\{|\downarrow\downarrow\rangle, |\uparrow\downarrow\rangle = c|\downarrow\downarrow\rangle, |\downarrow\uparrow\rangle = \hat{c}|\downarrow\downarrow\rangle, |\uparrow\uparrow\rangle = c\hat{c}|\downarrow\downarrow\rangle\}$. The extended BRST operator and ghost number are

$$Q = cH + \hat{c}\Pi, \quad N_g = cb + \hat{c}\hat{b} - 1. \quad (6.4)$$

The latter says that c, \hat{c} raise N_g by 1 while b, \hat{b} lower it by one, and the constant is chosen to make N_g anti-Hermitian. We take physical states to all lie at $N_g = 0$.

For help comparing to literature, we note that often only c, b are referred to as the ghosts and ghost momentum, while \hat{b}, \hat{c} respectively are the anti-ghosts and anti-ghost momentum. Additionally, an alternative convention is common in which (anti-)ghost momenta are defined as anti-Hermitian operators such as $\rho = ib$ (with anti-commutation relations $[c, \rho]_+ = i$). This has the advantage that no factors of i are required in the classical theory (e.g., for Poisson brackets as in footnote 23), but we choose our current notation as more likely to be familiar to most readers.

We also need to specify Hermiticity relations for the operators, which determine an (indefinite signature) inner product on the extended BRST state space. For this we will make choices that may seem rather strange at first sight [20, 46]. These will be necessary to end up with sensible Hermitian physical inner products, and in fact matches with standard prescriptions in gauge theory, albeit in a guise that may be unfamiliar. We will take the lapse N and its corresponding conjugate momentum Π to be *anti-Hermitian* operators.²⁶ Similarly we will take the ghosts \hat{c} and \hat{b} (corresponding to the constraint $\Pi = 0$) to be anti-Hermitian, so that the BRST charge Q is Hermitian. We retain the standard representation of $N, \Pi = -i\partial_N$ acting on wavefunctions $\psi(N)$ depending on $N \in \mathbb{R}$, but the anti-Hermiticity leads to a modified inner product with a sign flip:

$$\langle\psi_2|\psi_1\rangle = \int_{-\infty}^{\infty} \psi_2^*(-N)\psi_1(N)dN. \quad (6.5)$$

Notably, this inner product is not positive-definite: even wavefunctions have positive norm, while odd wavefunctions have negative norm. Nonetheless, just like for the ghost sector this is not a problem on the extended BRST state space, as long as we end up with a positive-definite physical inner product on BRST cohomology. Similarly, the inner product on the ghosts are fixed by their (anti-)Hermiticity. Of particular interest are the inner products between states with $N_g = 0$,

$$\langle\downarrow\uparrow|\downarrow\uparrow\rangle = \langle\uparrow\downarrow|\uparrow\downarrow\rangle = 0, \quad \langle\uparrow\downarrow|\downarrow\uparrow\rangle = \langle\downarrow\uparrow|\uparrow\downarrow\rangle = -1. \quad (6.6)$$

The Hermiticity relations imply that the value of $\langle\uparrow\downarrow|\downarrow\uparrow\rangle$ must be real, but the precise value is a normalisation convention (the sign is chosen for later convenience).

To connect this apparently strange choice of lapse quantisation to something more familiar, we recall some ideas from a covariant (e.g., Gupta-Bleuler) quantisation of gauge

²⁶An equivalent prescription is to redefine N, Π by a factor of i , which leads to Hermitian operators with wavefunctions defined on the imaginary axis. It is not entirely clear to us which is more natural.

fields. The (Hermitian) gauge fields A_μ are decomposed into annihilation/creation operators a_μ, a_μ^\dagger for each momentum, with covariant commutation relations $[a_\mu, a_\nu^\dagger] \propto \eta_{\mu\nu}$, and states are built on a vacuum annihilated by a_μ . This means that the timelike component (corresponding to A_0 , the Lagrange multiplier which imposes the Gauss law constraint) obeys a harmonic oscillator algebra with a wrong sign. If we try to write states in terms of wavefunctionals of A_0 , this leads to a vacuum state wavefunction with a wrong-sign Gaussian, $\psi \propto \exp(+\#A_0^2)$. To make sense of this (and match the inner products obtained from the oscillator representation), we can take this wavefunction to be defined for *imaginary* values of A_0 , so A_0 is a Hermitian operator with imaginary eigenvalues. In our conventions, N is analogous to iA_0 : an *anti*-Hermitian operator with real-valued wavefunctions. So, our anti-Hermitian lapse with its unusual inner product is essentially the same as the wrong sign appearing in timelike oscillators of gauge fields. For more details, see appendix F. The usual uniqueness of the representation of position and momentum operators on Hilbert space (the Stone-von Neumann theorem) does not apply because the inner product is not positive-definite.

Now we introduce some notation for the states in this BRST formalism which correspond to co-invariants and invariants as discussed in section 3:

$$|\psi_{\text{BRST}}\rangle\rangle := |\psi\rangle\rangle \otimes |\Pi = 0\rangle_N \otimes |\uparrow\downarrow\rangle_g, \quad (6.7)$$

$$|\psi_{\text{BRST}}\rangle := |\psi\rangle \otimes |N = 0\rangle_N \otimes |\downarrow\uparrow\rangle_g. \quad (6.8)$$

The first line gives co-invariant states corresponding to the BRST cohomology with ghost state fixed to be $|\uparrow\downarrow\rangle_g$ (annihilated by c but not by \hat{c}). To be BRST closed the state must be annihilated by Π , fixing the lapse sector state to be $|\Pi = 0\rangle_N$, with wavefunction $\langle N|\Pi = 0\rangle_N = 1$ (independent of N). BRST-exact states of this form are given by Q acting on $|\phi\rangle \otimes |\Pi = 0\rangle_N \otimes |\downarrow\downarrow\rangle_g$, which identifies $|\psi\rangle \sim |\psi\rangle + H|\phi\rangle$ recovering the co-invariant equivalence classes. The second line is a representation of invariant states by choosing ghost state $|\downarrow\uparrow\rangle_g$. To be BRST closed, such a state must be annihilated by H so we get invariants. The equivalence under adding BRST-exact states (co-invariance under the Π constraint) preserves only the integral $\int \psi(q, N) dN$ of the remaining wavefunction, and we have used this freedom to choose wavefunctions of the form $\psi(q)\delta(N)$, fixing the lapse state $|N = 0\rangle_N$.

6.4 Some examples

We are now ready to choose at some specific examples of gauge-fixing fermions Ψ , calculate the matrix elements of the associated BRST-trivial operator between (co-)invariant BRST closed states, and interpret the results.

6.4.1 Group averaging inner product

For our first example, choose the gauge-fixing fermion $\Psi = -ibN$ (with a factor of i required for Hermiticity since $N^\dagger = -N$). This gives us a BRST-exact operator $[Q, \Psi]_+ = -iNH + b\hat{c}$, which generates a finite gauge transformation $e^{[Q, \Psi]_+} = e^{-iNH}(1 + b\hat{c})$. We calculate matrix

elements of this operator between states $|\psi_{\text{BRST}}\rangle = |\psi\rangle \otimes |\Pi = 0\rangle \otimes |\uparrow\downarrow\rangle$ in the co-invariant sector:

$$\langle\langle \psi'_{\text{BRST}} | e^{[Q, \Psi]_+} | \psi_{\text{BRST}} \rangle\rangle = \langle\uparrow\downarrow | b\hat{c} | \uparrow\downarrow \rangle \int dN \langle \psi' | e^{-iNH} | \psi \rangle = \langle\langle \psi' | \psi \rangle\rangle_{\text{GA}}. \quad (6.9)$$

In the ghost sector, only the $b\hat{c}$ term contributes, giving us unity (interpreted as the trivial measure from fixing to synchronous gauge $\dot{N} = 0$). The integral over N comes from the inner product between $|\Pi = 0\rangle$ states. We end up with precisely the group average inner product (1.6) discussed in section 4.2.

To interpret the action of $[Q, \Psi]_+$ note that the bosonic part is the canonical Hamiltonian $\mathbf{H} = NH$, so we can interpret the result as evolution for unit coordinate time in the gauge $\dot{N} = 0$ as described in 6.2. The constant value of N can then be interpreted as the total proper time, which is a modulus that we must integrate over. The simple ghost part gives us the flat measure dN for this modulus [37].

To illustrate the invariance implied by the Fradkin-Vilkovisky theorem, we consider a few simple generalisations. First we can simply rescale the generator by inserting $e^{t[Q, \Psi]_+}$ for some real t , which has the interpretation of the coordinate time interval. The lapse integral from the bosonic piece (integrating over evolutions by proper time Nt) then results in an insertion of $2\pi\delta(tH)$, while the fermionic piece provides an extra factor of t : these almost precisely compensate for each other since $\delta(tH) = \frac{1}{|t|}\delta(H)$, together giving us a factor of $\text{sgn}(t)$. This illustrates the invariance under small deformations (changes of t when $t \neq 0$), but also the failure of invariance from discrete jumps that can occur when passing through a singular value: here, the change in sign as we pass through $t = 0$.

Another simple generalisation is to add an extra term $-\frac{\xi}{2}\hat{b}\Pi$ to the gauge-fixing fermion Ψ for some parameter ξ , which results in an extra Gaussian factor $e^{-\frac{\xi}{2}\Pi^2}$ in the inner product (without affecting ghost terms). This gives us Gaussian averaged gauges (often used in quantisation of gauge theories). For the lapse variable, the bosonic piece $-\frac{\xi}{2}\Pi^2 - iHN$ of $[Q, \Psi]_+$ can be written as a first-order differential operator acting on wavefunctions of Π . Solving the resulting time-dependent Schrödinger equation for the action of $e^{[Q, \Psi]_+}$ on initial lapse wavefunction $\delta(\Pi)$ gives us $e^{-\frac{\xi}{6}H^2}\delta(\Pi + H)$. So, ultimately we find that the insertion e^{-iNH} is replaced by $\exp\left(-\frac{\xi}{6}H^2 - iNH\right)$. The result after integrating over N is of course unchanged as expected, but the intermediate steps may be nicer, particularly in the semiclassical limit: the resulting integral over N will typically be sharply peaked at the classical value, giving a rapidly convergent integral rather than a conditionally convergent oscillatory integral.

We can also choose a much more general gauge for time evolution, where we fix coordinate time in essentially any way we please (e.g., conformal time in an FLRW universe rather than proper time). For this, consider a gauge-fixing fermion $\Psi = -ibNF$ where F is a positive Hermitian operator on the original unconstrained Hilbert space (a function of qs and ps). This gives us $[Q, \Psi]_+ = -ibcNFH - icbNHF + b\hat{c}F$. Classically (i.e., using Poisson brackets), the bosonic part of this is NFH , so we can interpret this as generating evolution using a rescaled Hamiltonian $\tilde{H} = FH$, along with a redefinition of the lapse N (given by $F^{-1}N$ in terms of the original lapse). The Π equation of motion sets this new rescaled lapse to be constant. The upshot is that we are choosing a gauge for coordinate

time t such that $\frac{d\tau}{dt} \propto F$ (where τ is the original proper time).²⁷ This is rather common in mini-superspace models (for example choosing $F(q)$ to make the target space metric flat, or using conformal time), and we will comment on this below.

Exponentiating this operator is slightly trickier because we have three terms which may not commute, but as we show in appendix C its matrix elements between $|\uparrow\downarrow\rangle$ ghost states relevant for co-invariants can be computed, finding

$$\langle\uparrow\downarrow|e^{[Q,\Psi]_+}|\uparrow\downarrow\rangle = F^{\frac{1}{2}} \exp\left(-iNF^{\frac{1}{2}}HF^{\frac{1}{2}}\right)F^{\frac{1}{2}}. \quad (6.10)$$

Taking the matrix elements of this result between $|\Pi = 0\rangle$ lapse states (i.e., integrating over N) corresponds to a group average with rescaled Hamiltonian $F^{\frac{1}{2}}HF^{\frac{1}{2}}$, with extra compensating ‘measure factors’ of $F^{\frac{1}{2}}$ (which perhaps could be absorbed into a redefinition of the wavefunctions and unconstrained inner product). By invariance under choice of Ψ , the result $F^{\frac{1}{2}}\delta(F^{\frac{1}{2}}HF^{\frac{1}{2}})F^{\frac{1}{2}} = \delta(H)$ is simply a new representation of the original group average rigging map, though the propagator in the integrand before averaging is different.²⁸

For sigma-models, if we take F to be a function $F(q) = \Omega(q)^{-2}$ of q only, then the rescaled Hamiltonian \tilde{H} is classically obtained by a conformal transformation of the target space metric $\tilde{G}_{ab} = \Omega^2 G_{ab}$ and a rescaled potential $\tilde{V} = \Omega^{-2}V$. Quantising including a ‘conformally coupled’ curvature term to the potential (as mentioned in 2.2) [25, 26], the action of the Hamiltonian is changed by a simple rescaling $\tilde{H} = \Omega^{-1-\frac{n}{2}}H\Omega^{\frac{n}{2}-1}$ (which is Hermitian under the inner product with rescaled measure $\sqrt{\tilde{G}}d^nq = \Omega^n\sqrt{G}d^nq$). This gives a simple correspondence between (co-)invariant wavefunctions under H or \tilde{H} : $\tilde{\psi}_{\text{inv}} = \Omega^{1-\frac{n}{2}}\psi_{\text{inv}}$ is annihilated by \tilde{H} if ψ_{inv} is annihilated by H , and similarly $\tilde{\psi}_{\text{co-inv}} = \Omega^{-1-\frac{n}{2}}\psi_{\text{co-inv}}$ induces a well-defined map between co-invariant cosets. Furthermore, taking a group averaging integral with \tilde{H} gives us essentially the same calculation as (6.10), up to conjugation by $\Omega^{\frac{n}{2}}$ (to account for the change in measure), and we arrive at²⁸ $\delta(\tilde{H}) = \Omega^{1-\frac{n}{2}}\delta(H)\Omega^{1+\frac{n}{2}}$. The upshot is that group averaging with a rescaled Hamiltonian \tilde{H} gives us the same physical Hilbert space, once we appropriately rescale wavefunctions.

6.4.2 Klein-Gordon form (and generalisations)

Next, consider choosing the gauge-fixing fermion $\Psi = -i\hat{b}\chi$ for some Hermitian operator χ on the unconstrained Hilbert space. For the Klein-Gordon form, we take χ to be given by a target space ‘time function’ $\chi(q)$ of the coordinates on \mathcal{M} , but as discussed in section 4 we can generalise this to an operator depending also on momentum. We find $[Q, \Psi]_+ = -i\Pi\chi + i[H, \chi]\hat{b}c$, noting the appearance of the ‘Fadeev Popov’ measure factor $i[H, \chi]$ in the ghost part. We now insert $e^{[Q,\Psi]_+}$ between two states to get an inner product (which

²⁷Alternatively, we could follow the approach sketched in section 6.2 by choosing $\Psi = -ibN - i\hat{b}F$, which gives $[Q, \Psi]_+ = -i(NH + \Pi F) + (\text{ghosts})$. Then, N is the usual lapse (not rescaled), but $[Q, \Psi]_+$ generates time translations in a gauge $\dot{N} = F$. This seems more natural, and makes closer contact with standard gauge-fixing approaches (see appendix F for $U(1)$ gauge theory in Lorenz gauge). Unfortunately, it is technically trickier to understand the resulting inner product.

²⁸A subtlety is that $F^{\frac{1}{2}}HF^{\frac{1}{2}}$ may not be essentially self-adjoint so the exponential may not be defined for all N , and hence $\delta(F^{\frac{1}{2}}HF^{\frac{1}{2}})$ may not make sense, or may be ambiguous: see comments in the discussion section 8.

is slightly trickier than for the group average because the two terms may not commute). We can write this exponential as

$$e^{[Q, \Psi]_+} = e^{-i\Pi\chi} + c\hat{b} [H, \Pi^{-1} e^{-i\Pi\chi}]. \quad (6.11)$$

One way to verify this is to replace $[Q, \Psi]_+$ by $s[Q, \Psi]_+$, and check that the result solves the ‘Schrödinger equation’ that the differentiating with respect to s gives the same as acting with $[Q, \Psi]_+$. We also explain a more explicit calculation in appendix C, which illustrates some general useful ideas for these sorts of calculations. Note that while $\Pi^{-1} e^{-i\Pi\chi}$ is singular at $\Pi = 0$, the singular part Π^{-1} does not contribute since it commutes with H .

The $c\hat{b}$ term will give us a non-zero result for matrix elements between ghost states $|\downarrow\uparrow\rangle$ corresponding to invariants. The lapse sector is easiest to evaluate in the momentum representation (wavefunctions of Π). Taking the simplest $|N = 0\rangle$ lapse wavefunctions ($\delta(N)$ in the N basis and a constant $\frac{1}{\sqrt{2\pi}}$ in the Π basis), we need the integral

$$\frac{1}{2\pi} \int d\Pi \frac{e^{-i\Pi\chi}}{\Pi + i\epsilon} = -i\Theta(\chi), \quad (6.12)$$

where the $i\epsilon$ in the denominator is an arbitrary prescription to regulate the $\Pi = 0$ singularity (other possible prescriptions differ by a constant which drops out when we take the commutator with H , for example the principle value integral gives $-i\pi \operatorname{sgn}(\chi)$). From this, we find the result

$$\llbracket \psi'_{\text{BRST}} | e^{[Q, i\hat{b}\chi]_+} | \psi_{\text{BRST}} \rrbracket = i \llbracket \psi' | [H, \Theta(\chi)]_- | \psi \rrbracket. \quad (6.13)$$

This is precisely the expression κ_Σ we found in (4.15) for the Klein-Gordon form on the surface Σ defined by $\chi = 0$.

We can also choose other lapse wavefunctions in the bra and ket states, which replaces the step function $\Theta(\chi)$ with a more general function which becomes zero or one for large negative or positive χ respectively. Specifically, this function is obtained by convolving the lapse wavefunctions with a step function. Thus we get the same result, but smoothed out over representations with shifted values of χ . A specific example is to use different fixed-lapse wavefunctions $|N = N_i\rangle$ and $|N = N_f\rangle$, which has the effect of shifting χ , replacing $\Theta(\chi)$ with $\Theta(\chi - N_f + N_i)$. The Gaussian average gauge (adding $-\frac{\xi}{2}\hat{b}\Pi$ to Ψ) similarly acts to smooth over a Gaussian window of width $\sqrt{\xi}$ in χ .

To connect this with the gauge-fixing κ map discussed in section 4, we simply take the inner product in the lapse and ghost sectors and treat the remaining object as an operator in the original unconstrained Hilbert space. A general Gaussian averaged gauge-fixing map looks like

$$\kappa_{\chi, \xi} = \langle N = 0 | \otimes \langle \downarrow\uparrow | e^{[Q, -i\hat{b}\chi - \frac{\xi}{2}\hat{b}\Pi]_+} | N = 0 \rangle \otimes | \downarrow\uparrow \rangle = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d\Pi}{\Pi} [H, e^{-i\Pi\chi}] e^{-\frac{\xi}{2}\Pi^2}. \quad (6.14)$$

6.4.3 Zero momentum gauge

For the above discussion, we took χ to be a general operator, with the specific example in mind of a function $\chi(q)$ of target space in a sigma-model to recover the Klein-Gordon form. Here we discuss a different choice which is a one-dimensional (or mini-superspace)

analogue to gauge-fixing on the maximal-volume slice in higher dimensions [15], mentioned in section 3.6.1.

For this, we choose $\chi = p_0$ where p_0 is the momentum in a timelike direction of target space. This is likely to correspond to a good gauge in cases where the potential V becomes large and negative for large q^0 . In that case, a typical trajectory will come from negative q^0 with negative p_0 (which corresponds to positive \dot{q}^0 due to the negative kinetic term), reflect off the potential, and return to negative q^0 with positive p_0 , crossing the $\chi = 0$ slice once in the positive direction.

An analogous choice in higher dimensional gravity is to gauge-fix to $K(x) = 0$, where K is the trace of the extrinsic curvature. This momentum corresponds to the timelike direction in superspace, and is canonically conjugate to the volume density $\sqrt{\det \gamma}$ where γ is the spatial metric. In most cases, the potential ($\sqrt{\det \gamma}(2\Lambda - R)$ where R is the spatial curvature, plus contributions from matter energy density) will become negative at large volume density for fixed conformal metric (the most notable exceptions occurring for positive cosmological constant). See [15] for a detailed discussion.

For simplicity we restrict our calculations to the example of a one-dimensional target-space discussed in section 4.5 with $H = -\frac{p^2}{2} + V(q)$ (now returning to the standard sign convention for p with $[q, p] = i$ so that $\dot{q} = -p$), and use the gauge-fixing function $\chi = p$. Our considerations straightforwardly extend to many more general cases: in particular, an immediate generalisation is a higher-dimensional target space whose metric G_{ab} is independent of q (so that p corresponds to a Killing vector of target space), in which case the following goes through unchanged by suppressing dependence on other coordinates.

To set expectations, recall that ignoring issues of operator ordering would give us a combination of a gauge-fixing delta function $\delta(p)$ with a measure factor $\Delta_{\text{FP}} = i[H, p] = -V'(q)$. But the two operators $\delta(p)$ and Δ_{FP} certainly do not commute, so it is not very obvious what the precise result will be.

As we saw above, the inner product comes from the $c\hat{b}$ term in $e^{i[Q, \Psi]_+}$ given in (6.11). We calculate the matrix elements in the position basis, with the only contribution to the commutator coming from the potential term in H :

$$\begin{aligned} \langle q|[V, \Pi^{-1}e^{-i\Pi p}]|q'\rangle &= \Pi^{-1} (\langle q|V(q)|q' + \Pi\rangle - \langle q - \Pi|V(q')|q'\rangle) \\ &= \delta(q' - q + \Pi) \frac{V(q) - V(q')}{\Pi}, \end{aligned} \quad (6.15)$$

where we use the fact that $e^{-i\Pi p}$ is the operator that translates q by Π . If we then integrate over Π (taking the expectation value with lapse wavefunctions $|N = 0\rangle$), we get matrix elements for the associated sesquilinear form

$$\kappa_p(q, q') = i\langle q|[H, \Theta(p)]|q'\rangle = -\frac{1}{2\pi} \frac{V(q) - V(q')}{q - q'}. \quad (6.16)$$

We can easily generalise this to the Gaussian averaged version as in (6.14),

$$\langle q|\kappa_{p, \xi}|q'\rangle = -\frac{1}{2\pi} \frac{V(q) - V(q')}{q - q'} e^{-\frac{\xi}{2}(q - q')^2}. \quad (6.17)$$

We can see that close to the diagonal $|q - q'| \ll 1$, these matrix elements become $\frac{1}{2\pi}$ times $-V'(q)$ which is the Fadeev-Popov determinant Δ_{FP} . This is a reasonable match with expectations because the position space matrix elements of $\delta(p)$ are simply constant, $\langle q|\delta(p)|q'\rangle = \frac{1}{2\pi}$.

Alternatively, we can calculate matrix elements in momentum space (if the potential does not grow too fast) to get

$$\langle p|\kappa_p|p'\rangle = \frac{1}{4\pi i}(\text{sgn } p - \text{sgn } p')\hat{V}(p - p'), \quad (6.18)$$

where $\hat{V}(p) = \int dq e^{-ipq}V(q)$, and take a Fourier transform to recover the position space result.

Note that these expressions for κ_p do not sharply project onto the $p = 0$ component of the wavefunction in any obvious sense. For example, we do *not* get the simplest possibility of an ‘average’ ordering of delta function and measure factor $\kappa = -\frac{1}{2}(\delta(p)V'(q) + V'(q)\delta(p))$. The way in which (6.16) (or better, its Gaussian average version (6.17)) gives a $p = 0$ gauge-fixing becomes sharper in a semi-classical limit, as we’ll see in section 7.

We can explicitly evaluate the overlap $\llbracket\psi_2|\psi_1\rrbracket_p$ of two invariant states using the sesquilinear form κ_p in various cases, and check that the result accords precisely with expectations from the interpretation in terms of gauge-fixing function $p = 0$. Here we just summarise results for various possible qualitative shapes of the potential.

First, in the motivating case where V is a positive constant for $q \rightarrow -\infty$ and negative at $q \rightarrow +\infty$, we find that $\llbracket\psi_2|\psi_1\rrbracket_p = \llbracket\psi_2|\psi_1\rrbracket_{\text{GA}}$, agreeing with the physical group average result. If instead V is a positive constant at $q \rightarrow +\infty$ and negative at $q \rightarrow -\infty$, we find a relative minus sign. This is because trajectories come in from the right (with negative \dot{q} and positive p) and leave from the left (with negative p), crossing the gauge-fixing slice in the decreasing direction. For a one-dimensional target space this is somewhat trivial (since the physical Hilbert space is one-dimensional), but this result also applies in more general cases where κ_p gives a useful representation of the physical inner product.

We can also check the case where V approaches a positive constant for both $q \rightarrow \pm\infty$, with scattering states as defined in (4.22). Similarly to (4.24) for the Klein-Gordon form, we can interpret the result most straightforwardly in terms of matrix elements between in and out states. We find

$$\llbracket\text{out}, \pm|\text{in}, \pm\rrbracket_p = \mp\llbracket\text{out}, \pm|\text{in}, \pm\rrbracket_{\text{GA}}, \quad \llbracket\text{out}, \pm|\text{in}, \mp\rrbracket_p = 0, \quad (6.19)$$

precisely in line with expectations from the gauge-fixing interpretation.

6.5 Mixing invariants and co-invariants

It is instructive to look not only at the inner products constructed as above, but also at the states obtained from acting with our gauge transformations $e^{[Q, \Psi]_+}$. For the case of group averaging $\Psi = -ibN$ acting on co-invariants, we find the following (where we represent the

matter and lapse sector as wavefunctions of q, N):

$$e^{[Q, \Psi]_+} \psi_0(q) | \uparrow \downarrow \rangle = \psi(q; N) (| \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle), \quad (6.20)$$

$$\text{where } i \frac{\partial}{\partial N} \psi(q, N) = H \psi(q, N), \quad \psi(q, N=0) = \psi_0(q). \quad (6.21)$$

So, we arrive at a wavefunction $\psi(q, N)$ which is a solution to the time-dependent Schrödinger equation, with N treated as the time!

The space of such wavefunctions gives us an alternative way to represent physical states. These are in a sense intermediate between invariants and co-invariants, because they are annihilated not by H or Π , but by a linear combination $H + \Pi$. Our gauge transformation has converted our original co-invariant wavefunction $\psi_0(q)$ to a new but physically equivalent form, namely the corresponding solution of the time-dependent Schrödinger equation $\psi(q, N)$. Furthermore, this new representation is less singular, in the sense that we can *directly* take inner products in the extended Hilbert space (remembering the antihermicity of N (6.5)) without encountering an indeterminate form:

$$\langle \psi | \psi \rangle = \int dq dN \psi^*(q, -N) \psi(q, N) = \frac{1}{2} \int dq dt \psi_0^*(q) e^{-iHt} \psi_0(q) \propto \langle \langle \psi_0 | \psi_0 \rangle \rangle_{\text{GA}}. \quad (6.22)$$

This inner product is finite, and unsurprisingly gives us the group averaging result (up to a factor of 2, which is compensated for by the ghosts, arising because this norm involves two insertion of $e^{[Q, \Psi]_+}$ so that $t = 2N$).

Just like (co-)invariants, this new representation of physical states also comes with an identification under addition of a null state in the image of H . Because these states are annihilated by $H + \Pi$, this is equivalent to addition of something in the image of Π , namely any total derivative with respect to N . In particular, this tells us that wavefunctions $\psi(q, N)$ and $\psi(q, N + \Delta N)$ related by a constant time translation ΔN are physically equivalent.

In terms of ghosts, we are choosing a representation of the BRST cohomology class which is not annihilated by c or \hat{c} , but by the non-hermitian combination $c - \hat{c}$. To make the resulting structure clearer, we can package the ghosts into a single complex (non-Hermitian) ghost $C \propto c - \hat{c}$ and its adjoint, and similarly package the two (anti-)Hermitian constraints into a single complex constraint G as follows:

$$G = \frac{H + \Pi}{\sqrt{2}}, \quad C = \frac{c - \hat{c}}{\sqrt{2}} \implies Q = C^\dagger G + C G^\dagger. \quad (6.23)$$

Now, a BRST closed state which is annihilated by C must be annihilated by G , and hence must be a solution of the time-dependent Schrödinger equation as above. Similarly, a BRST exact state which is annihilated by C must be of the null form described in the above paragraph. The key difference in the ghost sector (compared to states annihilated by c or \hat{c}) is that the state annihilated by C has non-zero norm.

Note that this is similar to the situation more typically encountered in BRST quantisation of gauge theories. In that case, one usually decomposes ghosts and constraints into positive frequency parts (analogous to C, G) and their conjugate negative frequency parts,

and builds states on a vacuum annihilated by positive frequency modes. See appendix F for details of the example of Maxwell theory. This means that we can immediately take inner products in the extended BRST Hilbert space. We expect a similar procedure to be useful for many problems in gravity (perturbation theory around Minkowski space being a simple example). However, it may not always be possible or convenient to pair up constraints in this manner. A simple example is in perturbation theory around an FLRW cosmology, where modes with non-zero momentum should be expected to pair up rather naturally in this way, but not spatially homogeneous zero modes, which may need to be treated separately perhaps using the ideas in this paper. Something similar occurs for the string worldsheet, where the analogues to G are the non-Hermitian Virasoro modes L_n for $n > 0$, with $L_n^\dagger = L_{-n}$, but the Hermitian zero modes L_0 are not paired in this way (see appendix E). We expect such issues to only become more pronounced when working around an excited background with fewer symmetries (e.g., a black hole spacetime).

7 Semiclassical limits

Our main examples of one-dimensional gravity are very useful to explore the main ideas without too much technical complication. But in some ways these examples are too simple: many calculations can be done exactly, which is rarely possible (perhaps even in principle) in realistic models for interesting applications. With this in mind we discuss the semiclassical limit in this section, aiming to illustrate ideas that generalise most readily to higher-dimensional theories of gravity.

We then illustrate these ideas concretely using the example of a one-dimensional target space (introduced in section 4.5) with $H = -\frac{p^2}{2} + V(q)$, emphasising the analogy with the maximal-volume gauge $\text{Tr } K = 0$ as discussed in section 6.4.3.

In this section we restore factors of \hbar to indicate how parameters scale in the semiclassical limit. Note that for gravity, the small parameter is $\hbar G_N$ (compared to a typical curvature scale), which we can scale separately from \hbar if we would like to use a semi-classical limit from gravity while treating matter as fully quantum.

7.1 General comments

We start our discussion with some general comments about the semi-classical limit in general sigma model or mini-superspace examples (2.4). This outlines a general strategy for understanding the Hilbert space in semi-classical perturbation theory, which we expect to generalise nicely to higher-dimensional theories of gravity. This is only a sketch, and we do not attempt to work out all the details (an interesting exercise for the future). In the next subsection we specialise to the examples of one-dimensional target space introduced in section 4.5, for which we implement these ideas explicitly.

7.1.1 Group averaging

We start by considering the group averaging rigging map and associated inner product. In the position basis (taking co-invariant states $|q\rangle\rangle$ represented by δ -function wavefunctions),

this is an integral of the propagator K (the position space matrix elements of $e^{-iHt/\hbar}$):

$$\eta(q, q') = \langle\langle q|q' \rangle\rangle = \int_{-\infty}^{\infty} dt K(q, q'; t), \quad K(q, q'; t) = \langle q|e^{-iHt/\hbar}|q' \rangle. \quad (7.1)$$

Note that this is similar to the fixed-energy retarded propagator (or Green's function) $G(q, q'; E) = i \int_0^{\infty} K(q, q'; t) e^{iEt/\hbar} dt$ at $E = 0$, the difference being that we integrate over all real t (not just $t > 0$). (In fact, it is twice the imaginary part of G , or the difference of propagators as expressed in (4.13).) This means we can use some standard results from semi-classical approximation in quantum mechanics (e.g., chapter 38 of [47]).

In the semi-classical limit, $K(q, q'; t)$ for fixed q, q', t is dominated by classical solutions beginning at q' and ending at q , taking time t . Each such solution contributes

$$K(q, q'; t) \sim \frac{1}{(2\pi i \hbar)^{\frac{n}{2}}} \sqrt{D_K(q, q'; t)} e^{\frac{i}{\hbar} S(q, q'; t)}, \quad (7.2)$$

where $S = \int dt(p\dot{q} - H(p, q))$ is the action of the classical solution (or Hamilton's principal function). We will not make use of the details of the prefactor here, except that D_K (the van Vleck determinant) is fixed in the $\hbar \rightarrow 0$ limit. For fixed q, q', t we generically expect there to be a discrete set of solutions — for given q' , there is an n dimensional space of solutions from choosing initial momenta, the same as the dimension of endpoints q — and the contributions from these are summed. Now to get the rigging map $\eta(q, q')$ we perform the integral over t , which is dominated by points of stationary phase. The time derivative of the on-shell action (with fixed endpoints) is the energy $\frac{\partial S}{\partial t} = -E$, so the stationary points correspond to classical solutions satisfying the constraints $H = 0$ (for any time t), each contributing

$$\eta(q, q') \sim \frac{1}{(2\pi i \hbar)^{\frac{n-1}{2}}} \sqrt{D_\eta(q, q')} e^{\frac{i}{\hbar} S(q, q'; E=0)} \quad (7.3)$$

with a different determinant D_η .²⁹ For q parametrically close to q' , there will also be a contribution from small t which needs to be considered separately.

7.1.2 Invariant (Wheeler-DeWitt) wavefunctions

Since η maps general wavefunctions to invariant states, we expect that $\psi(q) = \eta(q, q')$ in (7.3) should be a semi-classical solution to the Wheeler-DeWitt equation for any fixed q' . To leading order, this means that the exponent $S(q) = S(q, q'; E = 0)$ solves the Hamilton-Jacobi equation

$$H(q, p = \partial S(q)) = \frac{1}{2} G^{ab}(q) \partial_a S(q) \partial_b S(q) + V(q) = 0, \quad (7.4)$$

since at leading order the momentum $p_a = -i\hbar \partial_a$ acting on $e^{\frac{i}{\hbar} S(q)}$ (times any one-loop prefactor) can be replaced by $\partial_a S$. And indeed, all (real) solutions of this equation correspond

²⁹More generally, taking the Fourier transform of K with respect to t by stationary phase gives the Legendre transform of the action $S(q, q'; t)$ in the exponent, which is $S(q, q'; E) = \int p dq$ regarded as a function of E (Hamilton's characteristic function).

to the action of an $(n-1)$ -dimensional family of classical trajectories solving the constraint $E = 0$ (see [48], for example).

We can similarly get invariant states following the construction of scattering states in (3.8) or (3.9). In the case of (3.9) the semi-classical calculation is precisely the same as (7.3), except we only allow saddle-points at positive/negative t for in/out-states respectively. We then push the initial point q' to infinity (which means the saddle-points will be pushed to large $|t|$), and discard a divergent contribution to the action.

However, this does not make the closest possible contact with the classical theory, because $\psi(q) = \eta(q, q')$ does not represent a single classical solution: it corresponds to a superposition over all possible initial momenta compatible subject to the Hamiltonian constraint (or in the case of scattering states, a superposition over all impact parameters). To remedy this, we can consider the action of η on coherent wavefunctions with well-defined position and momentum in the $\hbar \rightarrow 0$ limit. For example, we can choose an initial Gaussian wavefunction of the form $\psi_0(q) \propto e^{-\frac{A}{2\hbar}(q-q_0)^2 + ip_0 q/\hbar}$ close to q_0 , where $H(q_0, p_0) = 0$.³⁰ Then $\psi(q) = \int \eta(q, q')\psi_0(q')$ will be supported close to the trajectory of the classical solution specified by initial data q_0, p_0 . The integral over q' has a real saddle point (namely, $q' = q_0$) only if the classical solution with initial conditions (q_0, p_0) passes through the final point q . Otherwise, the wavefunction is exponentially suppressed. In particular, this means that a pair of co-invariant states with Gaussian wavefunctions ψ_1, ψ_2 centered on phase space points (q_1, p_1) and (q_2, p_2) have an exponentially small inner product $\langle\langle \psi_2 | \psi_1 \rangle\rangle = \int dq_1 dq_2 \psi_2^*(q_2)\eta(q_2, q_1)\psi_1(q_1)$ unless (q_1, p_1) and (q_2, p_2) both lie on the same $E = 0$ classical solution. The statement that generalises to a higher dimensional theory of gravity is that the inner product is exponentially suppressed unless (q_1, p_1) and (q_2, p_2) correspond to initial data (metrics and extrinsic curvatures) for a pair of Cauchy surfaces Σ_1, Σ_2 in the same classical spacetime solving the Einstein equations. Such coherent states can also arise naturally from states prepared by Euclidean path integrals (e.g., the Hartle-Hawking no boundary state): see further comments in section 8.

Note that the exponential tails of the wavefunction away from the classical trajectory can be interesting, perhaps giving an amplitude to tunnel to a classically forbidden region (an example will be given in a moment). Furthermore, these non-perturbative parts of the wavefunction (or of an inner products between co-invariants) may be computed by a saddle-point for which either the trajectory q or the proper time/lapse t (or both) are complex, corresponding in higher dimensions to complex (or perhaps Euclidean) spacetime metrics. But importantly, the original definitions of the quantities of interest are in terms of integrals over real t and q , enabling us to determine which saddle-points contribute by appropriate deformation of the contour of integration. In particular, there is no conformal factor problem [49] in such cases where there is a clear Hilbert space interpretation of the quantity of interest, with a corresponding Lorentzian path integral (even if the relevant saddle-points may correspond to Euclidean geometries). Similar ideas have been suggested many times before [40, 50–52]. Note also that in such cases, contributing complex saddle

³⁰Here $A(q - q_0)^2$ should be interpreted as $A_{ij}(q^i - q_0^i)(q^j - q_0^j)$ for some positive-definite Riemannian metric on target space at the point q_0 (not the mixed-signature target space metric G appearing in H).

points can only give exponentially suppressed contributions, not enhanced [40]. We study an example in detail in section 7.2.6, and give further comments in section 8.

7.1.3 Gauge-fixing maps or inner products on invariants

It is relatively straightforward to apply the κ gauge-fixing maps to the semi-classical invariant wavefunctions described above. Starting with an invariant wavefunction which is well-localised along a single classical trajectory $(q_c(t), p_c(t))$, we should choose a gauge condition $\chi = 0$ which uniquely identifies a single point on the trajectory. For example, we might choose $\chi(q)$ to be a function of target space \mathcal{M} so that the curve $q_c(t)$ on which the wavefunction is large crosses the $\chi = 0$ surface exactly once. Then, applying the corresponding Klein-Gordon gauge-fixing operator $\kappa_\chi \sim \frac{1}{2}(p_\chi \delta(\chi) + \delta(\chi) p_\chi)$ gives us a wavefunction which is localised close to a single point on \mathcal{M} .

However, this is not convenient in the semiclassical limit because the resulting wavefunction $\kappa\psi$ is not well-localised in phase space: we have localised it *too much* in χ , so that it has large uncertainty in the conjugate momentum p_χ . For example, a purely ‘positive frequency’ state which corresponds to trajectories crossing the $\chi = 0$ slice with $\dot{\chi} > 0$ requires an exponentially delicate balance between $\delta(\chi)$ and $\delta'(\chi)$ terms in such a wavefunction. This issue is worse still for gauge conditions involving momenta, when the wavefunction does not localise on target space (the example $\chi = p$ is discussed below). We would prefer a gauge-fixing prescription which gives rise to co-invariant wavefunctions with well-defined position and momentum in the semi-classical limit, which allows an interpretation as classical initial data.

A resolution to this uses the Gaussian averaged ξ gauges introduced in section 6, where we smear the gauge-fixing delta function $\delta(\chi)$ to a Gaussian $\exp(-\frac{\chi^2}{2\hbar\xi})$. This localises χ to a width $\sqrt{\hbar\xi}$, and hence the conjugate momentum to a width of order $\sqrt{\hbar\xi^{-1}}$, meaning that the resulting wavefunction can correspond to a well-defined point in phase space.

To show how this works in more detail, for some $\chi(q, p)$ consider the action of $\kappa_{\chi, \xi}$ as written in (6.14) on a semiclassical invariant wavefunction. This is a wavefunction of the form $\psi(q) \sim e^{\frac{i}{\hbar}S(q)}$ (times a prefactor which is unaffected at one-loop order), where S solves the Hamilton-Jacobi equation (7.4). Note that S can be real (corresponding to a superposition of many classical solutions like in (7.3)), or have an imaginary part like our coherent wavefunctions introduced above in which case we can restrict attention close to the locus where $\text{Im } S$ is minimised.

The action of the operator $e^{-\frac{i}{\hbar}\Pi\chi}$ is computed by a path integral where we use χ as a ‘Hamiltonian’ and flow by ‘time’ Π . Classically, this means we evolve by $\frac{dq}{d\Pi} = \frac{\partial\chi}{\partial p}$, $\frac{dp}{d\Pi} = -\frac{\partial\chi}{\partial q}$, with initial condition $p(0) = S'(q(0))$ and final condition $q(\Pi) = q$ set by the point at which we evaluate the resulting wavefunction (in position basis). We will later integrate over Π weighted by $e^{-\frac{\xi}{2\hbar}\Pi^2}$, which will be dominated by small Π so we only need to expand the flow for short times, which is determined by local data in phase space. The new wavefunction is the exponential of the resulting on-shell action,³¹ which has small Π

³¹The one-loop determinant is $1 + O(\Pi)$, and does not contribute at the order we are working.

expansion $S(q) - \Pi\chi(q, \partial S(q)) + \frac{1}{2}\zeta(q)\Pi^2 + \dots$ where $\zeta(q) = \frac{\partial\chi}{\partial p_a} \left(\frac{\partial\chi}{\partial q^a} + \frac{\partial\chi}{\partial p_b} \frac{\partial^2 S}{\partial q^a \partial q^b} \right)$. This gives

$$\begin{aligned} [H, e^{-\frac{i}{\hbar}\Pi\chi}]e^{\frac{i}{\hbar}S(q)} &\sim H\left(q, \partial S(q) - \Pi\frac{\partial}{\partial q}\chi(q, \partial S(q)) + \dots\right)e^{\frac{i}{\hbar}(S(q) - \Pi\chi(q, \partial S(q)) + \frac{1}{2}\zeta(q)\Pi^2 + \dots)} \\ &\sim -\Pi\dot{\chi}(q)e^{\frac{i}{\hbar}(S(q) - \Pi\chi(q, \partial S(q)) + \frac{1}{2}\zeta(q)\Pi^2)}, \end{aligned} \quad (7.5)$$

where in the second line we use the Hamilton-Jacobi equation for H to identify $\frac{\partial H}{\partial p_a} \frac{\partial}{\partial q^a} \chi(q, \partial S(q)) = \frac{\partial\chi}{\partial q^a} \dot{q}^a + \frac{\partial\chi}{\partial p_a} \dot{p}_a = \dot{\chi}$ as the time derivative of χ along the classical trajectory with $p = \partial S$.

Now we perform the integral $-\frac{1}{2\pi\hbar} \int \frac{d\Pi}{\Pi} e^{-\frac{\xi}{2\hbar}\Pi^2}$ with the above result in the integrand as instructed in (6.14) (restoring a factor of \hbar from the time derivative). This is simply a Gaussian integral, as long as we can neglect higher-order terms in the expansion. This is self-consistent as long as the saddle-point value $\Pi = \frac{\chi}{\zeta + i\xi}$ is small (in fact, $\Pi^3 \ll \hbar$ to avoid contributions from higher-order terms at the one-loop order we are working). This means that q is close to the surface $\chi(q, \partial S(q)) = 0$ where the gauge condition is satisfied (or alternatively, for generic q if we choose $\xi \gg 1$). In that case, we find

$$\kappa_{\chi, \xi} e^{\frac{i}{\hbar}S(q)} \sim \frac{\dot{\chi}(q)}{\sqrt{2\pi\hbar(\xi - i\zeta)}} e^{\frac{i}{\hbar}S(q) - \frac{1}{2\hbar(\xi - i\zeta)}\chi(q, \partial S(q))^2}, \quad (1 \ll \xi \ll \hbar^{-1}). \quad (7.6)$$

Elsewhere, the wavefunction is exponentially suppressed. Thus, $\kappa_{\chi, \xi}$ imposes the gauge condition $\chi = 0$ not as a delta-function but as a Gaussian with width roughly $\sqrt{\xi\hbar} \ll 1$, accompanied with the classical value of the Fadeev-Popov factor $\dot{\chi}$. If we use $\kappa_{\chi, \xi}$ to compute the norm of the state, the integral $\int \psi^* \kappa_{\chi, \xi} \psi$ will be localised on the surface $\chi(q, \partial S(q)) = 0$.

The same calculation applies (using stationary phase approximation) for $\xi = 0$ when q is close to the $\chi = 0$ surface, but the wavefunction is no longer suppressed when $\chi \neq 0$. At points where $\chi(q, \partial S(q))$ is not small there can still be a real saddle-point computing $\kappa_{\chi} \psi(q)$, but it is no longer at small Π so it requires solving for the classical flow generated by χ for finite ‘time’. The Fadeev-Popov factor can then be more complicated and non-local in target space, which we have already seen for the case $\chi = p$ in (6.16) (explored more in the semiclassical limit below). Even when $\chi(q, \partial S)$ is small, there can be additional saddle-points at large Π which give further contributions to $\kappa_{\chi} \psi$ of comparable magnitude (which cannot happen for $\xi > 0$).

The resulting wavefunctions (7.6) can now be regarded as gauge-fixed representatives of co-invariant states, as outlined in section 3.6. But now we have a gauge fixing ansatz which applies for any gauge condition $\chi = 0$, and which behaves nicely in the semiclassical limit.

We can then use group averaging to compute norms of these states. First, we can check what happens locally when we group average the co-invariant state in (7.6), by concentrating on the region where $\frac{\chi^2}{\xi - i\zeta} \ll 1$ and small t . Time-evolution simply translates the Gaussian packet along the classical flow defined by S (with $p = \partial S$), so group averaging smears the wavefunction along each classical trajectory. We can change variable in this integral from t to χ to get a Gaussian integral (with a sign if $\dot{\chi} < 0$). The result is that we recover our original invariant wavefunction, up to a factor of $\text{sgn } \dot{\chi}$. This verifies that

$\kappa_{\chi,\xi}$ indeed acts as a generalised inverse to η , when $\chi = 0$ is a good gauge. But η will give the correct (positive) physical result in all cases. In particular, η may have additional contributions from large times in the group average integral: for example, these may arise if χ is a good gauge locally or perturbatively, but fails to be a good gauge globally or with non-perturbative effects (with trajectories crossing the $\chi = 0$ surface several times).

7.1.4 Comparison to symplectic quotient phase space

At this level of approximation, we expect that the resulting Hilbert space to be equivalent to the quantisation of the symplectic quotient phase space (i.e., the phase space of classical zero-energy trajectories). We sketch one strategy for making this identification here.

It is simplest to quantise a phase space if it can be identified as the cotangent bundle $T^*\mathcal{Q}$ of some configuration space \mathcal{Q} , so we can take the Hilbert space $L^2(\mathcal{Q})$ to consist of wavefunctions defined on \mathcal{Q} . But there is no general way to do this for the symplectic quotient. An exception is for constraints which are linear functions of momenta. In that case, the gauge transformations generated by the flow act on configuration space (for us, the target space \mathcal{M}), since δq is a function of q^a only. The symplectic quotient can then be taken at the level of configuration space: it is canonically identified with $T^*(\mathcal{M}/\mathbb{R})$ (where \mathbb{R} is our one-dimensional gauge group generated by a single constraint). Of course, this does not apply to our gravitational constraints. However, we can reduce to such a situation if we limit ourselves to working perturbatively around a family of background solutions. For this, we first choose an $(n-1)$ -dimensional family of trajectories. This sweeps out an n -dimensional surface in phase space, which we can (locally, and at generic points) write as $p = p_0(q)$. We then work with a constraint to linear order in $\delta p = p - p_0(q)$, namely $H \sim \frac{\partial H}{\partial p_a}(q, p_0(q))\delta p_a$, so our perturbative symplectic quotient is given by the cotangent space of \mathcal{M} , modulo the flow generated by the vector field $v^a = \frac{\partial H}{\partial p_a}(q, p_0(q))$ (the velocity \dot{q}^a of the background trajectory). We can make this more concrete by choosing a gauge condition $\chi(q, p) = 0$ which selects a single point along each background trajectory, so that the physical phase space is identified as the cotangent bundle of the $(n-1)$ -dimensional submanifold $\chi(q, p_0(q)) = 0$ of \mathcal{M} .

To implement this quantum mechanically, we start with a ‘background’ semi-classical wavefunction $\psi_0(q) = A(q)e^{\frac{i}{\hbar}S(q)}$ which solves the Wheeler-DeWitt equation (so $S(q)$ solves the Hamilton-Jacobi equation and A contains one-loop and higher order corrections). This encodes the $(n-1)$ -dimensional family of classical trajectories described above, with $p_0(q) = \partial S(q)$. Then we take an ansatz for invariant wavefunctions of the form $\psi(q)A(q)e^{\frac{i}{\hbar}S(q)}$, where $\psi(q)$ is held fixed in the $\hbar \rightarrow 0$ limit. Requiring that this wavefunction solves the Wheeler-DeWitt equation to leading nontrivial order tells us that $v^a \frac{\partial \psi}{\partial q^a} = 0$, where $v^a = \frac{\partial H}{\partial p_a}(q, \partial S)$ is the tangent vector to gauge transformations as above (this is equivalent to $\delta S = -i\hbar \log \psi$ solving the Hamilton-Jacobi equation perturbed around the background solution S). So as expected from the above, our states are wavefunctions on $\mathcal{Q} = \mathcal{M}/\mathbb{R}$, where \mathbb{R} is the group generated by the flow of v^a . And if $\chi(q, p)$ is a good gauge-fixing function then solutions to $v^a \partial_a \psi = 0$ are in one-to-one correspondence with initial data giving ψ on the $(n-1)$ -dimensional surface $\chi(q, \partial S) = 0$ (at least locally near $\chi = 0$).

To check the identification of our Hilbert space with $L^2(\mathcal{Q})$, we also need to compare inner products, which we can compute as matrix elements of $\kappa_{\chi,\xi}$. Using (7.6), at one loop order we indeed get an inner product on the $(n-1)$ -dimensional surface $\chi(q, \partial S) = 0$ of the form $\int_{\chi=0} |\psi|^2$ (the choice of measure can be absorbed into a change of A in the background state ψ_0 , with a corresponding rescaling of ψ). Alternatively, we can treat $\kappa_{\chi,\xi}(\psi A e^{\frac{i}{\hbar} S})$ as a gauge-fixed representative of a co-invariant state (as outlined in section 3.6), and use group averaging to compute the inner product (dominated by short times or small lapse). We expect these to give the same results to all orders in perturbation theory if χ is a good gauge for our background trajectories (but group averaging should be taken as the correct physical prescription for going beyond perturbation theory).

Note, however, that this simple Hilbert space of functions with L^2 inner product applies only at this order. The formalism we have developed can be used to systematically compute corrections to the states and inner product. The loop corrections will typically contain higher and higher derivatives, indicating that summing corrections leads to an inner product which is non-local in configuration space.

7.2 One-dimensional target examples

To illustrate these ideas more concretely, we return to the example of a one-dimensional target space (see section 4.5) with $H = -\frac{p^2}{2} + V(q)$.

In particular, we mostly consider the case where the potential V approaches a positive constant $\frac{\omega^2}{2}$ for $q \rightarrow -\infty$ and monotonically decreases, becoming negative for $q > q_0$ where q_0 is the unique zero of V . Then there is a unique classical solution (solving the constraints, and up to gauge transformations), which comes in from $q \rightarrow -\infty$, bounces off the potential at $q = q_0$, and returns back to $q \rightarrow -\infty$. As already remarked in sections 3.6.1 and 6.4.3, this could be a simple model for an expanding and re-collapsing bang/crunch cosmology, or a model for different bulk time slices in an asymptotically AdS spacetime with the same asymptotic boundary time. With the latter in mind, many of our remarks will be closely analogous to the discussion in [15]. The physics in this case is somewhat trivial (with a one-dimensional Hilbert space) but it is nonetheless sufficient to demonstrate many ideas that readily generalise, and we will remark on such generalisations along the way.

7.2.1 Group averaging

We begin by considering the group average inner product $\langle\langle q|q' \rangle\rangle$. Each trajectory contributes to the semi-classical inner product (7.3) as

$$\langle\langle q|q' \rangle\rangle \sim \frac{1}{\sqrt{\dot{q}_i \dot{q}_f}} \exp\left(\frac{i}{\hbar} \int_{q'}^q p dq\right) + \dots, \quad (7.7)$$

where $p = \pm\sqrt{2V(q)}$ in the integrand, and the one-loop prefactor contains the initial and final velocities $\dot{q}_{i,f}$ of the classical solution.³² For fixed q' , this is a WKB solution to $H|\psi\rangle = 0$ as a function of q (and the same holds with $q \leftrightarrow q'$).

³²Some care is required to keep track of extra phases from the turning point $q = q_0$ where the solution reflects and \dot{q} changes sign. For example, this leads to the $\frac{\pi}{4}$ phase shift familiar from WKB solutions.

Now, in the main case of interest where $q, q' < q_0$ both lie within the classically allowed region $V < 0$ (and $q \neq q'$), there are in fact four zero-energy classical solutions which contribute. One trajectory goes directly from q' to q and one reflects from the potential barrier at q_0 , and also we have the time-reverse of each of these. Summing the contributions from all four solutions gives a product of WKB scattering wavefunctions $\langle\langle q|q' \rangle\rangle = \psi_{\text{WKB}}^*(q')\psi_{\text{WKB}}(q)$, indicating that in this case the Hilbert space is one-dimensional. In contrast, if we had a potential which is everywhere positive there would be two distinct classical solutions (left-moving and right-moving) and the group average becomes a sum of two terms $\eta(q, q') = \psi_+^*(q')\psi_+(q) + \psi_-^*(q')\psi_-(q)$.

In section 7.2.6, we will return to calculation of $\eta(q, q')$ in cases where there is no classical solution from q' to q (one endpoint lies in the classically forbidden region, or the endpoints are separated by a potential barrier).

7.2.2 Invariant (scattering) wavefunctions

In the case when $V(q) < 0$ for $q > q_0$, there is a single solution to the Wheeler-DeWitt equation which does not blow up as $q \rightarrow \infty$. This is the WKB wavefunction $\psi_{\text{WKB}}(q)$ which appeared in the calculation of η above. Nonetheless, there are two natural basis vectors for this one-dimensional Hilbert space — the scattering in state and out state — and the relative phase contains physical information. Here we compute the wavefunction of an in-state (an out-state is obtained by complex conjugation). This is defined as in 3.8 by a long evolution of a momentum eigenstate $e^{-iHT}|p\rangle$ in the limit $T \rightarrow \infty$ (choosing $p = -\omega$ to solve the constraint in the asymptotic ‘free’ region $q \rightarrow -\infty$, and normalising similarly to (4.22)). We could immediately write down the answer as a WKB wavefunction, but instead we will take the opportunity to illustrate methods which generalise more readily.

Anticipating that we will compute inner products of invariant states using the $p = 0$ gauge-fixing, we will initially compute the wavefunction in the momentum basis, and restrict to evaluating at $p = 0$. That is, we compute the matrix element $\langle p = 0 | e^{-iHT} | p = -\omega \rangle$ in the large T limit. We can compute this by a (configuration space) path integral on the interval $t \in [-T, 0]$ (holding the final time fixed to get a limiting classical solution $q_c(t)$ when we take $T \rightarrow \infty$ at fixed t). The initial and final states $|p\rangle$ are specified by Neumann boundary conditions $\dot{q}(-T) = \omega$, $\dot{q}(0) = 0$. The corresponding action is

$$S = \int_{-T}^0 \left(-\frac{1}{2} \dot{q}^2 - V(q) \right) dt - q(-T)\omega, \quad (7.8)$$

where we have added a boundary term at $t = -T$ to get a good variational problem with the specified boundary conditions.

A semi-classical calculation begins by looking for a classical solution $q_c(t)$, here with boundary conditions $\dot{q}_c(0) = 0$ and $\dot{q}_c(-T) = \omega$. The solution for $T \rightarrow \infty$ has $q_c(0)$ at the turning point q_0 (with $V(q_0) = 0$) to give zero energy, so the initial point $q(-T) \sim -\omega T$ in the far asymptotic region has velocity ω . Explicitly, this solves $\frac{1}{2}\dot{q}_c^2 = V(q_c)$ so $\int_{q_c(t)}^{q_0} \frac{dq}{\sqrt{2V(q)}} = |t|$. The on-shell action is $-\int_{q_c(-T)}^{q_0} \sqrt{2V(q)} dq - \omega q_c(-T)$, which in the $T \rightarrow \infty$ limit becomes

$$S_{\text{in}} = - \int_{-\infty}^{q_0} (\sqrt{2V(q)} - \omega) dq - q_0\omega. \quad (7.9)$$

This is the WKB phase at the turning point of a zero energy scattering state.

Loop corrections are accounted for by expanding fluctuations around the classical solution, $q = q_c + \delta$. At one loop we need only the quadratic action $-\frac{1}{2} \int_{-\infty}^0 (\dot{\delta}^2 + V''(q_c(t))\delta^2) dt$, and we integrate over $\delta(t)$ with initial and final boundary conditions $\dot{\delta} = 0$ (since we are working with states of fixed momentum). This action is that of harmonic oscillator with time-dependent frequency (and wrong sign kinetic term), with Hamiltonian $H_{\text{one-loop}} = -\frac{p^2}{2} + \frac{1}{2}V''(q_c(t))\delta^2$. The corresponding time-dependent Schrödinger equation is solved by

$$\psi_{\text{one-loop}}(\delta, t) = \frac{1}{\sqrt{f(t)}} \exp\left(\frac{1}{2i\hbar} \frac{\dot{f}(t)}{f(t)} \delta^2\right), \quad (7.10)$$

where $\dot{f}(t) = V''(q_c(t))f(t)$. By differentiating the classical equation of motion $\ddot{q}_c = V'(q_c)$, one solution is $f = \dot{q}_c$, and this gives the correct initial constant wavefunction (zero momentum for δ , normalised as $\frac{1}{\sqrt{\omega}}$) as $t \rightarrow -\infty$. The $\dot{\delta}(0) = 0$ boundary condition is implemented by taking the overlap with the final zero-momentum wavefunction at time $t = 0$ (i.e., integrating over δ with a factor $(2\pi\hbar)^{-1/2}$), which gives us the result for the one-loop determinant $\frac{1}{\sqrt{i\dot{f}(0)}} = \frac{1}{\sqrt{iV'(q_0)}}$. So, we get the one-loop result

$$\psi_{\text{in}}(p = 0) \sim \frac{\exp(\frac{i}{\hbar} S_{\text{in}})}{\sqrt{iV'(q_0)}} \quad (7.11)$$

for the momentum-space wavefunction evaluated at $p = 0$.

Our semi-classical calculation generalises straightforwardly to compute the momentum-space in-state wavefunction at different values of p with $|p| < \omega$. The action (7.8) gets an additional boundary term $-pq(0)$, and we look for classical solutions with $\dot{q}(0) = -p$. We use the same solution up to a time-translation to obey the boundary condition, so that $V(q(0)) = \frac{p^2}{2}$. The on-shell action is similar to (7.9) with upper integration limit given by the endpoint $q(0)$, and the addition of the boundary term $-pq(0)$. For $p > 0$ we need to integrate from $-\infty$ to the turning point q_0 , and then back to the endpoint after going to the other branch of the square root (with $\dot{q} < 0$) to follow the reflected trajectory. For small momentum, the resulting action is

$$S(p) = S_{\text{in}} - pq_0 - \frac{p^3}{6V'(q_0)} + \dots \quad (7.12)$$

We can check this result by comparing with the WKB approximation, which gives the position space wavefunction

$$\psi_{\text{in}}(q) \sim e^{\frac{i}{\hbar} S_{\text{in}} + i\frac{\pi}{4}} \frac{2 \cos\left(\frac{1}{\hbar} \int_q^{q_0} \sqrt{2V} - \frac{\pi}{4}\right)}{(2V(q))^{\frac{1}{4}}} \quad (7.13)$$

for $q < q_0$. This is normalised to give $\psi_{\text{in}} \sim \frac{1}{\sqrt{\omega}} e^{-\frac{i}{\hbar} \omega q} + \dots$ at $q \rightarrow -\infty$, where \dots indicates the reflected part $\propto e^{+\frac{i}{\hbar} \omega q}$. Near the turning point $q = q_0$, this becomes the Airy function

$$\psi_{\text{in}}(q) \sim e^{\frac{i}{\hbar} S_{\text{in}}} \sqrt{\frac{2\pi\hbar}{iV'(q_0)}} \lambda \text{Ai}(\lambda(q - q_0)), \quad (7.14)$$

where $\lambda^3 = \frac{-2V'(q_0)}{\hbar^2}$. This is precisely the Fourier transform of our momentum space result (7.12) at small p , including the correct phase and one-loop factor from (7.11).

We can also directly calculate the position-space wavefunction $\psi_{\text{in}}(q) \propto \lim_{T \rightarrow \infty} \langle q | e^{-iHT} | p \rangle$ similarly, by imposing a Dirichlet boundary condition fixing $q(0)$. In the classically-allowed region $q < q_0$ there will be two classical solutions contributing: one coming directly from infinity and one reflected from the potential (again, these always come from the same underlying solution up to time-translation). These contribute the two exponentials in the cosine in (7.13). The one-loop calculation is similar, except the different boundary conditions mean that we evaluate the one-loop wavefunction (7.10) at $\delta = 0$ to give the usual WKB denominator proportional to $(\dot{q})^{-1/2} = (\pm 2V(q))^{-1/4}$.

For $p^2 > \omega^2$ or for $q > q_0$ there are no real classical solutions. A semi-classical analysis in that case requires complex solutions (which also can contribute important non-perturbative corrections to the calculations we have already considered). We discuss this in section 7.2.6.

7.2.3 Zero momentum gauge-fixing

We now examine the sesquilinear form κ_p obtained in section 6.4.3 from the gauge-fixing fermion $\Psi = -i\hat{b}p$, which we expect to give us a physical inner product, interpreted as coming from imposing the gauge condition $p = 0$. We found the explicit position-space expression (here resotting \hbar)

$$\llbracket \psi_2 | \psi_1 \rrbracket_p = -\frac{1}{2\pi\hbar} \int dq_1 dq_2 \psi_2^*(q_2) \frac{V(q_1) - V(q_2)}{q_1 - q_2} \psi(q_1). \quad (7.15)$$

This does not localise to the $p = 0$ component of the states $\psi_{1,2}$ in any obvious way. Here we see how to nonetheless evaluate it from the $p = 0$ value of the invariant states in a semi-classical limit, order-by-order perturbation theory.

The crucial observation is that (7.15) will be dominated by the vicinity of the classical turning point q_0 where $V(q_0) = 0$ in the semi-classical limit. For the classically allowed region $q < q_0$ the integrals over $q_{1,2}$ will be highly oscillatory, while for the classically forbidden region $q > q_0$ the wavefunction will be exponentially small. Zooming in on this region, the factor $\frac{V(q_1) - V(q_2)}{q_1 - q_2}$ in the vicinity of this turning point becomes simply a constant $V'(q_0)$ to leading order, and we get

$$\llbracket \psi_2 | \psi_1 \rrbracket_p \sim (-V'(q_0)) \left(\int_{-\infty}^{\infty} \frac{dq_2}{\sqrt{2\pi\hbar}} \psi_2^*(q_2) \right) \left(\int_{-\infty}^{\infty} \frac{dq_1}{\sqrt{2\pi\hbar}} \psi(q_1) \right). \quad (7.16)$$

The integrals in the brackets give precisely the momentum-space wavefunctions evaluated at $p = 0$, which we write as $\psi_{1,2}(p = 0)$. So this one-loop order result is given by evaluating the invariant wavefunctions on the gauge-fixing surface $p = 0$, along with the classical value of the Fadeev-Popov determinant (with operator ordering ambiguities not arising at this order in \hbar). In a system with more degrees of freedom, we would also be left with an integral over the unfixed transverse variables.

We can use this to evaluate the norm of the in-state $\llbracket \psi_{\text{in}} | \psi_{\text{in}} \rrbracket_p$, using the result (7.11) for the $p = 0$ wavefunction: the measure factor of $-V'(q_0) > 0$ cancels the one-loop determinant to get unit norm, agreeing with the result from group averaging as computed in section

4.5. We can also compute the overlap between in-state and out-state (with momentum-space wavefunction $\psi_{\text{out}}(p) = \psi_{\text{in}}^*(-p)$) obtaining $\llbracket \psi_{\text{out}} | \psi_{\text{in}} \rrbracket_p \sim i \exp(\frac{2i}{\hbar} S_{\text{in}})$, which is the correct scattering phase to order \hbar^0 .

7.2.4 Higher orders in perturbation theory

Now we would like to illustrate the strategy for extending this to higher order in perturbation theory. The philosophy here is that we are given only the zero momentum component $\psi^{p=0}$ of the wavefunctions in a semi-classical expansion (perhaps by extending the above analysis to higher loops), and we would like to compute the inner product only in terms of this data. For this, the first step is to Taylor expand $\frac{V(q_1) - V(q_2)}{q_1 - q_2}$ around the turning point:

$$\frac{V(q_1) - V(q_2)}{q_1 - q_2} \sim \sum_{n=1}^{\infty} \frac{V^{(n)}(q_0)}{n!} \frac{(q_1 - q_0)^n - (q_2 - q_0)^n}{q_1 - q_2} = V'(q_0) + \frac{1}{2} V''(q_0)(q_1 + q_2 - 2q_0) + \dots \quad (7.17)$$

The n th term in this expansion will contribute at n -loop order in perturbation theory, with insertions of powers $(q - q_0)^{n-1}$ with $q = q_1$ or q_2 . This gives us integrals of the form $\int (q - q_0)^k \psi(q) dq$, which is not determined only by the momentum wavefunction at $p = 0$, but by its k th derivative there $\psi^{(k)}(p = 0)$. These derivatives are not part of the given data, but we can nonetheless obtain them by using the fact that our wavefunctions solve the Wheeler-DeWitt equation, perturbatively near $q = q_0$ (or at small p in momentum space). Importantly, we need only local data (derivatives of the potential at $q = q_0$).

Let's see how to do this explicitly, choosing a solution normalised so that $\psi^{p=0} = \frac{1}{\sqrt{2\pi\hbar}}$ (so $\int_{-\infty}^{\infty} \psi(q) dq = 1$). Expanding near the turning-point, the wavefunction will be given by an Airy function (or we can use the semi-classical momentum space result (7.12)):

$$\psi(q) \sim \lambda \text{Ai}(\lambda(q - q_0)), \quad \psi(p) \sim \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{p^3}{3\lambda^3} - \frac{i}{\hbar} p q_0} \quad (7.18)$$

where $\lambda^3 = \frac{-2V'(q_0)}{\hbar^2}$. This suggests writing the Wheeler-DeWitt equation in terms of the variable $u = \lambda(q - q_0)$ and expanding for large λ with u held fixed, which gives

$$\frac{d^2 \psi}{du^2} - u \psi(u) = \sum_{n=2}^{\infty} \lambda^{1-n} \frac{V^{(n)}(q_0)}{n! V'(q_0)} u^n \psi(u) = \frac{1}{\lambda} \frac{V''(q_0)}{2V'(q_0)} u^2 \psi(u) + \dots \quad (7.19)$$

Now if we make the perturbative ansatz $\psi(u) \sim \sum_{n=0}^{\infty} \lambda^{1-n} \psi_n(u)$ with $\psi_0(u) = \text{Ai}(u)$, at fixed order in λ we will have an equation giving $\psi_n''(u) - u \psi_n(u)$ in terms of lower-order ψ_k with $k < n$. This will have a unique solution once we fix $\int \psi_n(u) du = 0$ (one coefficient in the two-dimensional space of solutions is fixed by decaying at $u \rightarrow \infty$, and the remaining coefficient of $\text{Ai}(u)$ is fixed by the integral condition fixing the $p = 0$ wavefunction³³). To first order, we get

$$\psi_1(u) = \frac{V''(q_0)}{10V'(q_0)} (u^2 \text{Ai}'(u) - u \text{Ai}(u)) = \frac{V''(q_0)}{10V'(q_0)} (\text{Ai}^{(5)}(u) - 5 \text{Ai}''(u)). \quad (7.20)$$

³³These integrals may not converge due to growing rapidly oscillating tails as $u \rightarrow -\infty$, but they are nonetheless well-defined in a distributional sense (defined by the Fourier transform evaluated at $p = 0$, for example). No such issue appears if we work directly in momentum space.

Higher order solutions can all be similarly written as sums of derivatives of $\text{Ai}(u)$. In fact, it is rather nicer to do this calculation in momentum space (and more directly compatible with semi-classical calculations); we do this systematically in appendix D.1.

The upshot is that the required derivatives $\psi^{(k)}(p=0)$ are determined order-by-order in perturbation theory in terms of the original data $\psi(p=0)$ by solving the Wheeler-DeWitt equation, in terms of local data (derivatives of the potential at q_0). Using this data in the expansion of (7.17) in turn determines the inner product $\llbracket \psi_2 | \psi_1 \rrbracket_p$, from only the data $\psi_{1,2}(p=0)$ and a finite number of derivatives of V at the turning point at each order in perturbation theory. We explain how to systematically carry this out in appendix D.1, finding a first correction at order \hbar^2 :

$$\frac{\llbracket \psi_2 | \psi_1 \rrbracket_p}{\psi_2^*(p=0)\psi_1(p=0)} = -V'(q_0) \left(1 - \frac{3V''(q_0)^3 + V^{(4)}(q_0)V'(q_0)^2 - 4V^{(3)}(q_0)V'(q_0)V''(q_0)}{12V'(q_0)^4} \hbar^2 + \dots \right).$$

We also give two examples in the appendix, including one (an exponential potential) with no perturbative corrections to the leading $-V'(q_0)$ result at any order in \hbar , but nonetheless with non-perturbative corrections.

7.2.5 Gaussian average gauge fixing

Above, we have used $p=0$ gauge-fixing to evaluate inner products of invariant states. We can also make use of this gauge fixing as the map κ_p to construct gauge-fixed co-invariant representatives of a state (as discussed in sections 1.2, 3.6). However, κ_p does not result in a state which is well-localised in position space: it results in a wavefunction proportional to $\frac{V(q)}{q-q_0}$ (to leading order in \hbar). To get a state which is well-localised near the turning point in position space (as well as near $p=0$ in momentum space), we can instead use the Gaussian averaged version $\kappa_{p,\xi}$ of the gauge-fixing map.

We evaluate $\kappa_{p,\xi}$ acting on the in-state semi-classical wavefunction (in position space) using the result (6.17). This results in the integral

$$\kappa_{p,\xi}\psi_{\text{in}}(q) \sim -\frac{e^{\frac{i}{\hbar}S_{\text{in}}+i\frac{\pi}{4}}}{2\pi\hbar} \int dq' \frac{V(q)-V(q')}{q-q'} \frac{1}{(-2V(q'))^{\frac{1}{4}}} e^{-\frac{\xi}{2\hbar}(q-q')^2 - \frac{1}{\hbar} \int_{q_0}^{q'} \sqrt{-2V(x)} dx}, \quad (7.21)$$

where we have written the state so that the action is manifestly real in the classically forbidden region $q' > 0$. For $q' < 0$ we should include both possible branches $q' + i\epsilon$ (the incoming part) and $q' - i\epsilon$ (the reflected part).

We now evaluate this by saddle-point, with the saddle point equation giving $\xi(q-q') = -\sqrt{-2V(q')}$. For q close to q_0 (specifically, $|q-q'| \ll \xi^{-2}$), we find a saddle in the classically forbidden region at $q' - q_0 \sim -\frac{\xi^2}{2V'(q_0)}(q-q_0)^2 > 0$, with action $-\frac{1}{2}\xi(q-q_0)^2$. So performing the one-loop integral around this saddle gives us a Gaussian wavefunction

$$\kappa_{p,\xi}\psi_{\text{in}}(q) \sim e^{\frac{i}{\hbar}S_{\text{in}}+i\frac{\pi}{4}} \sqrt{\frac{-V'(q_0)}{2\pi\hbar}} e^{-\frac{\xi}{2\hbar}(q-q_0)^2} \quad (7.22)$$

close to its peak. To verify this, we can integrate against ψ_{in}^* (which is simplest if we first Fourier transform to momentum space so we avoid needing the Airy function) to recover the norm $\llbracket \psi_{\text{in}} | \psi_{\text{in}} \rrbracket$.

This does not agree with our general result in (7.6). The reason is that the semi-classical form of the position space wavefunction breaks down at the turning point q_0 where the gauge condition $p = 0$ is satisfied. However, it does precisely agree if we instead work in momentum space.

7.2.6 Non-perturbative effects

We finish our discussion of the semi-classical approximation with examples of interesting non-perturbative effects in the group average inner product. These also serve as simple examples where our strategy of imposing constraints as operator in a quantum theory gives different results to the alternative of imposing constraints classically before quantising.

To illustrate the ideas in a simple setting, we will study the model with quadratic potential $V(q) = \frac{1}{2}\lambda^2(q^2 - a^2)$. This has two distinct classically allowed regions $q < -a$ and $q > a$, separated by the barrier $|q| < a$ where $V(q) < 0$. We will be interested in contributions to the group average inner product which come from tunnelling under this barrier.

First, we compute the propagator $\langle q|e^{-iHt/\hbar}|q'\rangle$. For any endpoints q, q' and any time t , there is a unique classical solution to $\ddot{q} = \lambda^2 q$ with $q(0) = q'$ and $q(t) = q$. The on-shell action of this solution is

$$S(t) = \frac{1}{2}a^2\lambda^2 t - \lambda \frac{(q^2 + q'^2) \cosh(\lambda t) - 2q'q}{2 \sinh(\lambda t)}, \quad (7.23)$$

and including a one-loop determinant gives us the semi-classical approximation for the propagator (which happens to be exact for the quadratic potential):

$$\langle q|e^{-\frac{i}{\hbar}Ht}|q'\rangle = \sqrt{\frac{\lambda}{-2\pi i \hbar \sinh(\lambda t)}} e^{\frac{i}{\hbar}S(t)}. \quad (7.24)$$

However, these are not generally solutions of the full gravitational theory because they do not solve the constraint $H = 0$. This is imposed when we group average, since the integral over t is dominated semi-classically by saddle-points where $S'(t) = -E = 0$. The energy is

$$E(t) = -S'(t) = -\frac{1}{2}a^2\lambda^2 - \lambda^2 \frac{(q^2 + q'^2) - 2 \cosh(\lambda t) q'q}{2 \sinh^2(\lambda t)}. \quad (7.25)$$

Now if $q, q' < -a$ (so both endpoints lie in the same classically allowed region), there are four real solutions to $E(t) = 0$. If we write $q = -a \cosh(\lambda t_1)$ and $q' = -a \cosh(\lambda t_2)$ (so that $t_{1,2}$ are the times for $E = 0$ solutions to get from q, q' to the turning point), then these real solutions are $t = \pm t_1 \pm t_2$, where the signs can be chosen independently.

However, in other cases there are no real solutions: for example, when $q = -a \cosh(\lambda t_L) < -a$ and $q' = a \cosh(\lambda t_R) > a$ (for $t_L, t_R > 0$) so the endpoints are separated by the potential barrier. Nonetheless, we have the group average inner product written as an integral over real t , which correspond to off-shell Lorentzian geometries (off-shell because they do not solve the constraint, which is the g_{tt} equation of motion from varying the lapse), and this integral is nonzero. We can compute it by deforming the real t defining contour into the complex plane.

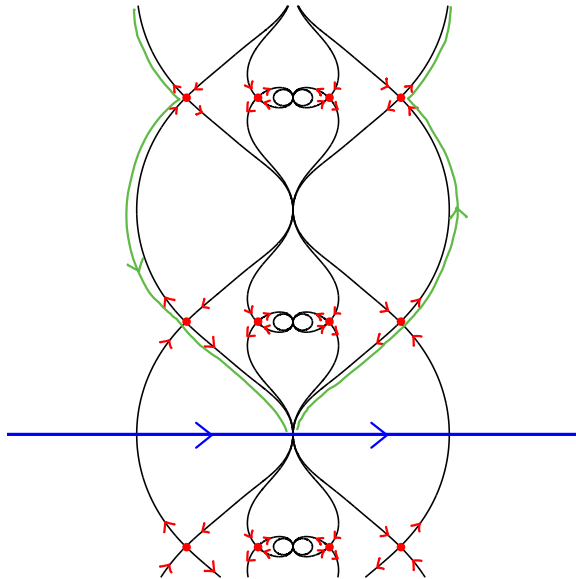


Figure 4. The complex t plane for the group averaging integral computing $\langle\langle q|q' \rangle\rangle$, where q, q' are separated by a potential barrier. Red dots indicate saddle points $E = 0$. Steepest descent contours are shown in black, with the red arrows indicating the descent direction (i.e., the direction of increasing $\text{Im} S$). The figure is periodically repeated in the imaginary direction. The original real t contour of integrating is in blue. We can deform this to the green contour, since $\text{Im} S$ is large and positive as $\text{Im} t \rightarrow +\infty$.

In this parameterisation, the energy can be written as

$$E(t) = -a^2 \lambda^2 \frac{(\cosh(\lambda t) + \cosh(\lambda(t_L + t_R)))(\cosh(\lambda t) + \cosh(\lambda(t_L - t_R)))}{2 \sinh^2(\lambda t)}. \quad (7.26)$$

From this we see that there are four families of solutions to $E(t) = 0$, namely $t = \pm t_L \pm t_R + \frac{i\pi}{\lambda}(1 + 2n)$, where the signs are chosen independently and n is an integer. Importantly, the group averaging interpretation of our path integral means that we know the defining Lorentzian (real t) contour,³⁴ so we can identify the correct way to deform the integration contour to a path of steepest descent, and identify which of these complex saddle-points contribute (and with what sign or phase). The appropriate deformation is shown in figure 4.

The integrand is suppressed (i.e., S has positive imaginary part) in the upper half-plane $\text{Im} t > 0$, so we deform the contour in that direction and find that the appropriate contour passes through the saddle-points at $t = \pm(t_L + t_R) + \frac{i\pi}{\lambda}(1 + 2n)$ for $n \geq 0$. We can interpret the $n = 0$ solutions as a trajectory which travels to the turning point in Lorentzian time (real lapse), tunnels under the barrier with a Euclidean geometry (imaginary lapse) taking imaginary time $i\frac{\pi}{\lambda}$, and then continues to the final point with another real Lorentzian evolution. Alternatively we can choose an equivalent complex path (for example with constant

³⁴In our case, this integral over real lapse (after doing all other integrals) is convergent. In other cases (such as target spaces with sufficiently high dimension) this may not be true because of the singularity at $t = 0$, and more care is required to identify the correct prescription. A closely related point is discussed in [53].

complex lapse) and deformations can be thought of as complex-valued diffeomorphisms: the invariant definition of a solution is a homotopy class of paths on the Riemann surface $H(q, p) = 0$ with complex q, p .

The integral is dominated by the two saddle points at $t = \frac{i\pi}{\lambda} \pm (t_1 + t_2)$, which gives us

$$\langle\langle q|q' \rangle\rangle \sim \frac{2}{a\lambda\sqrt{\sinh(\lambda t_L)\sinh(\lambda t_R)}} e^{-\frac{1}{2\hbar}a^2\pi\lambda} \cos\left(\frac{a^2\lambda}{4\hbar}(\sinh(2\lambda t_L) - 2\lambda t_L + \sinh(2\lambda t_R) - 2\lambda t_R)\right). \quad (7.27)$$

In particular, we do not have contributions from the saddle points at $t = \frac{i\pi}{\lambda} \pm (t_1 - t_2)$, which are equally dominant.

Even in the case when there are real saddle-points, there can be additional non-perturbative contributions from complex saddle-points with important physical implications. In the case $q = -a \cosh(\lambda t_1) < -a$ and $q' = -a \cosh(\lambda t_2) < -a$, a similar analysis shows that in addition to four real saddle-points at $t = \pm t_1 \pm t_2$, the steepest descent contour passes through complex saddle-points at $t = \pm(t_1 + t_2) + \frac{2\pi i}{\lambda}n$ for $n = 1, 2, 3, \dots$. Despite providing exponentially small corrections (scaling as $e^{-\frac{1}{\hbar}a^2\pi\lambda}$), they have a qualitatively important effect. The real saddles give a result which factorises as $\langle\langle q|q' \rangle\rangle \sim \psi(q)\psi^*(q')$ to all orders in perturbation theory, indicating that all the states $|q\rangle$ for $q < a$ are physically equivalent up to normalisation. But the leading non-perturbative contributions violate this factorisation, indicating that the physical Hilbert space is in fact two-dimensional (since they describe mixing with the classical solution on the other side of the barrier).

An important and general feature is that we do not have any contribution from the paths with t in the lower half-plane, which have $\text{Im } S < 0$ and hence would give exponentially large results. It is clear that such saddles can never lie on the contour of steepest descent, since the defining integral runs over a contour with oscillatory integrand, so the integral must be suppressed in the limit $\hbar \rightarrow 0$ (never enhanced).

An equivalent way to express the fact that the contributing saddles all lie in the upper half-plane $\text{Im } t \geq 0$ is that the lapse has (or rather, can be chosen to have) non-negative imaginary part $\text{Im } N \geq 0$. This is a very simple version of the Kontsevich-Segal-Witten criterion for allowable complex metrics [54, 55]. This raises a question of whether such a result holds more generally: if we begin with a Lorentzian gravitational path integral and deform to its the steepest-descent contour, does any saddle-point along the path necessarily obey this criterion? This seems necessary for there to be a consistent perturbation theory around the relevant complex spacetimes.

Note that if we impose the constraints classically in this example, we eliminate these interesting non-perturbative effects entirely. The symplectic quotient phase space consists of two points, corresponding to the solutions on the left or right side of the barrier. Quantising this gives a two-dimensional Hilbert space, but there is no sense in which the states on the left and right can mix.

Non-perturbative effects can have a yet more dramatic effect, even changing the dimension of Hilbert space. For example, take a cubic potential $V(q) = -\lambda(q - q_1)(q - q_2)(q - q_3)$ with $\lambda > 0$ and three real roots $q_1 < q_2 < q_3$. Now the classically allowed region consists of $q < q_1$, as well as the finite interval $q_2 < q < q_3$. In the latter region, there is a classically

bound orbit giving a second point in the constrained phase space. So constraining first would lead to a two-dimensional Hilbert space. But non-perturbative effects render this bound state metastable, and in fact the Hilbert space has a single state. It would be an interesting exercise to understand the details of this from a semi-classical perspective.

More generally (especially in realistic non-renormalisable theories of gravity with a complicated perturbation theory), one might worry that exponentially small non-perturbative corrections are negligible and should not be taken seriously if we do not completely understand the high-order perturbative behaviour. The examples in this section illustrate a more general principle of why they are nonetheless informative: the interesting non-perturbative contributions are of a qualitatively different nature. In particular, perturbative corrections to the gravitational inner product are all local in superspace (and come from the small lapse/shift part of the group average integral). The non-perturbative corrections can be completely non-local, because they involve a finite amount of time evolution. This allows them to manifest phenomena which are invisible in perturbation theory, such as tunnelling in the examples above, or (as commented on in section 8.3) non-local contributions to the inner product in black hole interiors inducing null states

8 Discussion

8.1 Additional considerations in higher dimensions

More constraints

The most obvious complication in higher dimensions is that we must deal with many constraints: we now have a Hamiltonian constraint $\mathcal{H}(x)$ for every spatial point x , and we also have momentum constraints $\mathcal{P}_i(x)$. There are correspondingly many choices in how we account for the constraints in the quantum theory. We can define physical wavefunctions as invariants (solutions to the Wheeler-DeWitt equation), co-invariant equivalence classes, or a mixture. There is probably no universal best strategy: it is useful to have this freedom in mind to make whatever choice is most convenient or illuminating for a given problem of interest.

In the context of perturbation theory around a static background (such as Minkowski space), we can use a quantisation scheme which makes close contact with approaches familiar from gauge theory (or the string worldsheet). This involves dividing the constraints into (non-Hermitian) positive frequency parts and their conjugate negative frequency parts with respect to time translation (under a conveniently gauge-fixed Hamiltonian). We then choose physical states which are annihilated by the positive frequency parts (invariance), and impose equivalence relations under negative frequency parts (co-invariance). Typically, these respectively remove unphysical transverse and longitudinal polarisations of photons/gluons/gravitons. In BRST language, this means choosing representatives of cohomology classes which are annihilated by the positive frequency part of ghost operators. See appendix F for details of this in abelian gauge theory. The subtleties of inner products that have been the main subject of this paper are absent in this case, since we impose invariance under non-Hermitian combinations of constraints as in the simple example in

section 6.5. Nonetheless, particularly for closed universes there can be zero modes which do not divide into positive/negative frequency parts and we must deal with these constraints separately.

More generally, for perturbations around time-dependent backgrounds (e.g., for cosmology or black holes) there is no obvious choice of positive-frequency modes, so other approaches are likely to be useful. For example, for perturbation theory around homogeneous cosmological spacetimes it may be useful to use the methods of this paper to treat the zero modes of constraints, while finding a positive/negative frequency split for spatially inhomogeneous perturbations.

Field-dependent structure constants

The additional challenges of constraints in more realistic models of gravity is not primarily from their number, but rather from their nature. Their algebra is not only non-abelian, but also has field-dependent structure constants. Specifically, the Poisson bracket of two Hamiltonian constraints depends on the inverse spatial metric γ^{ij} :

$$\left\{ \int N_1 \mathcal{H}, \int N_2 \mathcal{H} \right\} = \int \gamma^{ij} (N_1 \partial_j N_2 - N_2 \partial_j N_1) \mathcal{P}_i. \quad (8.1)$$

This leads to additional technical complications, for example when defining measures for group averaging. The usual Fadeev-Popov procedure does not immediately work because the field-dependence means that gauge transformations do not form an ordinary group.

Fortunately, the BFV/BRST formalism takes care of these generalisations rather elegantly. For gravity in particular, the BRST charge continues to take the form familiar from gauge theories, namely c times constraints plus a term with bcc contracted into structure constants (for more general algebras there can be terms with more ghosts, $b^2 c^3$ etc. [23]):

$$Q = \int d^d x \left[c \mathcal{H} + c_{\perp}^i \mathcal{P}_i - i b_{\perp}^{\perp} c_{\perp}^j \partial_j c_{\perp}^i + i b c_{\perp}^i \partial_i c + i \gamma^{ij} b_{\perp}^{\perp} c \partial_j c \right]. \quad (8.2)$$

This is valid classically (with $\{Q, Q\} = 0$); for the quantum BRST charge we have to resolve operator ordering in a way which depends on details of the quantisation. With Q in hand, we can continue to define physical states as the cohomology, pure gauge operators as BRST exact commutators $[\Psi, Q]$, and so forth (adding also $\int (\hat{c} \Pi + \hat{c}_{\perp}^i \Pi_{\perp}^i)$ for the primary constraints in analogy to (6.4), where Π and Π_{\perp}^i are the momenta conjugate to lapse N and shift N_{\perp}^i respectively).

For example, to compute a group averaged inner product (in synchronous gauge $\dot{N} = \dot{N}_{\perp}^i = 0$) we might use the gauge-fixing fermion $\Psi = -i \int (bN + b_{\perp}^{\perp} N_{\perp}^i)$, generalising section 6.4.1. With this, $\{Q, \Psi\}$ has bosonic part $-\int (N \mathcal{H} + N_{\perp}^i \mathcal{P}_i)$, generating a diffeomorphism.³⁵ But the ghost part will contain a term of the form $\gamma^{ij} N b_{\perp}^{\perp} \partial_j c$ from the last term in the BRST charge: this gives us a measure for the group average integral which depends on γ^{ij} .

³⁵In the quantum theory, ordering constants required by Hermiticity of Q might also give rise to shifts in $\mathcal{H}, \mathcal{P}_i$. This is similar to the case of non-unimodular Lie groups, where the left-invariant and right-invariant Haar measures differ so the correct measure for the group average is not so obvious, and the Dirac/Wheeler-DeWitt condition for invariant states needs to be modified [33]. This is automatically incorporated in the BRST formalism, where demanding Hermiticity of Q requires a quantum shift in the constraint term [20].

Another example of the complications of field dependent structure constants arises in the maximal-volume gauge-fixing $\text{Tr } K = 0$. The corresponding pure gauge generator $\{Q, \int \hat{b} \text{Tr } K\}$ contains a term proportional to $\int \gamma^{ij} \hat{b} b_i^\perp c \partial_j c$, quadratic in ghosts. Presumably this will be necessary to retain gauge-invariance in the ideas of [15] beyond one loop in perturbation theory.

Backgrounds with symmetries

We are often interested in backgrounds which preserve some symmetry group G , and perturbation theory around such a background comes with additional subtleties. Classically, there are linearisation stability constraints [56]. This means that not every solution to the Einstein equation linearised around the background corresponds to a family of exact solutions: there is a constraint that perturbations should have vanishing charges which generate G .

In the quantum theory, the $G_N \rightarrow 0$ limit should not simply reduce to QFT on a fixed spacetime, but we should also gauge by G (since the isometries are the diffeomorphism symmetries which are not spontaneously broken by the background). This naturally shows up when we calculate an inner product via group averaging, since the group average includes an integral over the isometries, and these cannot be accounted for by any gauge-fixing in the symmetric background. The result is that after taking gauge fixing into account, at $G_N \rightarrow 0$ we are still left with a group average integral over G . A prominent example is de Sitter space [8, 14, 42, 57].

8.2 Singularities: constraints are not essentially self-adjoint

There is one feature of the gravitational constraints (appearing even for mini-superspace models) that we have not addressed at all so far: they typically fail to be essentially self-adjoint. In simple terms, they act as Hermitian operators on sufficiently nice (e.g., smooth, compactly supported) functions, but do not have a unique definition on a larger domain, usually because some boundary conditions need to be specified. While this may seem to be mostly a technical issue (preventing operators from being exponentiated to finite gauge transformations, for example), it points at an important physical property of gravity.

In fact, this quantum phenomenon is a symptom of a classical illness. The Cauchy problem in gravity does not allow us to evolve for an arbitrary amount of time (in a given gauge) for general initial data. Instead, we often encounter a singularity in finite time. These can be curvature singularities, or they can be coordinate singularities (occurring even in Minkowski space for a poor choice of initial Cauchy surface and/or gauge).

In the classical context, one usually deals with this simply by terminating spacetime and its evolution at such a singularity, professing ignorance of what might lie beyond. In the symplectic quotient construction of the classical phase space, we declare equivalence between two sets of initial data only if they correspond to data on a pair of Cauchy surfaces in a common singularity-free (but not necessarily geodesically complete) spacetime. One could guess at additional equivalences that occur because of evolution through a singularity to some new classical spacetime, but that would be extremely speculative. Equally, we do

not exclude a point from the phase space if its evolution leads to a singularity either in the future (such as a black hole) or in the past (such as a big bang).

We advocate that this conservative approach can guide an analogous prescription for the quantum theory, at least semi-classically. In the path integral (for example, computing inner products by group averaging), this would simply mean integrating over smooth spacetimes, which in the semi-classical limit amounts to allowing only smooth metrics as saddle-points (though making this precise may lead to yet more subtleties when determining whether a given complex saddle-point contributes). Note that this restriction only applies to the part of spacetime (in a finite band of time) relevant to the saddle-point in question; singularities in an extension the classical solution further into the future or past are allowed (excluding these, as in the similar prescription of [58], would project out certain states which we regard as physical). In operator terms, the upshot is that we do not choose a self-adjoint extension; instead we accept that gauge transformations no longer generate unitaries, roughly because any part of the wavefunction that evolves to a singularity gets discarded. We spell out the details in the example of JT gravity [5]. In this simple solvable 2D model, this approach is perfectly satisfactory even non-perturbatively, giving a well-defined Hilbert space whose classical limit agrees with the standard phase space.

We can contrast our approach with an alternative suggested in DeWitt’s treatise [?], which postulates a boundary condition that invariant wavefunctions should vanish in limits of singular metrics. But this amounts to a reflecting boundary condition (generalising Dirichlet boundary conditions for a non-relativistic particle on a half-line), asserting that gravitational collapse to a singularity precedes a ‘bounce’ back to a re-expanding classical geometry (at least locally; presumably spatial gradients make the actual dynamics rather complicated). For example, in DeWitt’s discussion of a Friedmann cosmology this boundary condition turns a classical bang-crunch spacetime into a cyclic ‘bouncing’ cosmology, where every crunch is followed by a new big bang. We regard this idea that evolution continues through a singularity and re-emerges to a classical spacetime as highly speculative.

8.3 Null states

One of our main motivations for revisiting canonical methods in quantum gravity arose from the phenomenon of ‘null states’ induced by non-perturbative gravitational effects. It has long been appreciated that the perturbative Hilbert space of gravity (around a fixed classical background) can appear to contain too many states, in extreme kinematic regimes. The archetypal example is the interior of an old black hole (after time of order $(\hbar G_N)^{-1}$), which can allow more than $\exp(\frac{A}{4\hbar G_N})$ states, in tension with a microscopic state-counting interpretation of Bekenstein-Hawking entropy. A possible resolution is that the states constructed in perturbation theory are not in fact linearly independent, so that various states are in fact ‘null’ — they have zero norm, and the physical states are taken to be equivalence classes under such wavefunctions. This requires modifications to the physical inner product which are non-perturbatively small (i.e., exponential in $(\hbar G_N)^{-1}$) for individual matrix elements, but finely tuned. Various recent results from the path integral [2, 31, 59, 60] give indirect and direct evidence that non-perturbative gravitational

effects can supply precisely the needed corrections to the inner product.³⁶

However, from this path integral perspective these null states seem rather miraculous: for a matrix of inner products to have a large kernel (but no negative eigenvalues), its entries must be finely tuned in a carefully coordinated way. Such a coincidence demands an explanation. We propose that the underlying mechanism is gauge symmetry, with null states arising from the usual redundancy under gauge transformations. But we are not invoking some new gauge symmetry: instead, we attribute the null states to the usual diffeomorphism symmetry of gravity.³⁷ More specifically, we refer to the non-perturbative part of diffeomorphisms which are not close to the identity.

We can make this more concrete in the language of co-invariants. A perturbative Hilbert space might be constructed by choosing a gauge-fixing which is good to all orders in perturbation theory. For example, if we have a gauge condition which selects a unique Cauchy surface in the dominant semi-classical spacetime, we can use this to make an ansatz for states as sketched in section 3.6 (perhaps using a κ map to construct the ansatz and compute an inner product order-by-order in perturbation theory). But even after this gauge-fixing, there could be additional equivalences arising from a residual group of unfixed diffeomorphisms which are far from the identity. Geometrically, this might happen when there are spacetimes contributing to the inner product (perhaps encoding non-perturbative effects as in section 7.2.6) which contain several Cauchy surfaces satisfying the gauge condition.

This idea can be made completely precise in a simple toy model [60] where the non-perturbative part of diffeomorphism symmetry is the symmetric group $\text{Sym}(n)$ permuting several components of space. We hope that the ideas of this paper might be useful to implement a similar idea in more realistic theories of gravity. Additionally, this canonical perspective (which does not rely on any asymptotic boundary or AdS/CFT-type duality) may help to understand other circumstances under which similar effects are relevant, for example in cosmological spacetimes.

8.4 Topology change

Having motivated the interest of non-perturbative gravitational effects, we should point out that the most interesting examples involve topology change. Such effects are very natural in the context of path integrals where the relevant spacetimes have Euclidean signature. However, topology change is incompatible with ordinary Lorentzian geometry, since any

³⁶At least, gravitational effects can provide some of these modifications, with sums over geometries being sufficient to make gravity compatible with the Bekenstein-Hawking entropy. But there remains the outstanding issue of the factorization problem [61, 62], whereby exchange of closed ‘baby’ universes leads to superselection sectors for long-distance physics. As yet, there does not seem to be any evidence that gravity alone is sufficient to solve this without deliberate fine-tuning, so perhaps stringy or other effects are necessary.

³⁷We may need to allow a mild extension of the usual diffeomorphism, for example to permit jump discontinuities to account for topology-changing effects (see [60] and the next subsection). More speculatively, perhaps diffeomorphisms are part of an extended (perhaps stringy) gauge symmetry which is restored in the UV, which may be behind additional null states such as those required for factorization (see footnote 36).

two Cauchy surfaces (in a globally hyperbolic spacetime) must have the same topology. Given this, one might reasonably believe that these effects are impossible to capture in a canonical formalism like we have explored in this paper, since this typically leads to an interpretation in terms of Lorentzian spacetimes.

Fortunately, this is too pessimistic since topology change is possible if we allow a mild relaxation of a Lorentzian geometry, where the metric is permitted to have certain degenerations along surfaces of codimension two (or higher). One example is the Louko-Sorkin ‘crotch singularity’ [63], which locally looks like an n -fold cover of Minkowski space ramified over a codimension-2 branch locus. Such singularities appears in replica wormhole calculations [1, 2] from the Lorentzian perspective emphasised in [64–66], and similar configurations have been used to calculate the entropy of black holes [52] and de Sitter space [67] from Lorentzian path integrals.

But while these Lorentzian path integrals make closer contact with the perspective of this paper, we still lack a complete understanding of topology change from an *ab initio* canonical perspective. This might look like ‘interaction terms’ in the Hamiltonian (perhaps required for a consistent self-adjoint completion of the constraints) which couple sectors of different spatial topology. We believe that filling this gap is currently the most pressing problem in canonical quantum gravity.

Acknowledgements We would like to thank Shoaib Akhtar, Steve Carlip, Don Marolf and Mukund Rangamani for many helpful discussion and comments on drafts. JH is supported by U.S. Department of Defense through the National Defense Science and Engineering Graduate (NDSEG) Fellowship Program as well as NSF grant PHY-2408110 and funds from the University of California. HM is supported by DOE grant DE-SC0021085 and a Bloch fellowship from Q-FARM. This research was supported in part by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP).

A Proof of Klein-Gordon causal propagator identity

Let ψ be a solution to the constraint $H\psi = 0$ with compactly supported initial data, and f a source function with compact support in time. Write $\psi_+ = \mathcal{G}_+ f$ for the retarded solution to $H\psi_+ = f$. We have the identity

$$\frac{1}{2}\nabla \cdot [\psi_+ \nabla \psi - \psi \nabla \psi_+] = f\psi, \quad (\text{A.1})$$

using $H = -\frac{1}{2}\nabla^2 + V$. Now, since f is compactly supported in time we can choose a Cauchy surface Σ lying to the future of the support of f . Let \mathcal{M}_- denote the region to the past of Σ . Since the right hand side of our identity is supported only on \mathcal{M}_- , we can evaluate $\int_{\mathcal{M}} f\psi$ by integrating the left hand side over \mathcal{M}_- . This integral evaluates to a boundary term on Σ only (there is no boundary term in the past, because ψ_+ is the retarded solution and vanishes there):

$$\begin{aligned} \int_{\mathcal{M}} f\psi &= \frac{1}{2} \int_{\Sigma} (\psi_+ \nabla_N \psi - \psi \nabla_N \psi_+) \\ &= \Omega(\psi, \Delta f), \end{aligned} \quad (\text{A.2})$$

where for the second line we use the fact that $\psi_+ = \Delta f$ on Σ , since it lies to the future of the support of f .

B Static target spaces

Here we consider the group-averaging and Klein-Gordon inner products in the special case of a static target space. This means that \mathcal{M} is symmetric under translations and reversals of a ‘time’ coordinate $T = q^0$, with the remaining coordinates q^i denoted by X^i ; we have $G_{00} = -f(X)$ for a positive function f , $G_{0,i} = 0$, and $G_{ij}(X)$ a Euclidean metric on a manifold Σ . For simplicity of exposition we will additionally assume that Σ is compact, and take the potential $V(X)$ to be positive (for example, a constant $\frac{m^2}{2}$ with positive nonzero ‘mass’ m). Then the Hamiltonian is

$$H = -\frac{1}{2}f(X)p_T^2 + \mathcal{D}, \quad (\text{B.1})$$

where \mathcal{D} is a strictly positive Hermitian operator acting on X variables only (namely, the potential V minus half the Laplacian on \mathcal{M} acting on functions independent of T).

Now, we can decompose a general wavefunction on \mathcal{M} into simultaneous eigenfunctions of H and p_T ,

$$f(T, X) = \frac{1}{\sqrt{2\pi}} \sum_n \int d\omega \hat{f}_n(\omega) e^{-iT\omega} \phi_n^\omega(X), \quad (\text{B.2})$$

$$(\mathcal{D} - \frac{1}{2}\omega^2 f(X))\phi_n^\omega(X) = E_n(\omega)\phi_n^\omega(X). \quad (\text{B.3})$$

Here, ϕ_n^ω are a complete set of eigenfunctions for the operator $\mathcal{D} - \frac{1}{2}\omega^2 f$, which (for any fixed ω) has a discrete spectrum by our assumption that Σ is compact (relaxing this, we could replace n by continuous labels, such as ‘spatial momenta’ k in the case that \mathcal{M} is Minkowski space with constant $V = \frac{m^2}{2}$). We normalise them such that $\int_\Sigma dX \sqrt{-G} (\phi_m^\omega(X))^* \phi_n^\omega(X) = \delta_{mn}$.

The energy eigenvalues $E_n(\omega)$ are even functions of ω (since they depend only on ω^2 with $E_n(\omega) > 0$, and strictly decreasing for $\omega > 0$). This last property follows from ‘time-independent perturbation theory’, which gives E'_n as an ‘expectation value’ of f :

$$E'_n(\omega) = -\omega \int_\Sigma dX \sqrt{-G} f(X) |\phi_n^\omega(X)|^2. \quad (\text{B.4})$$

In particular, from this it follows that $E_n(\omega) = 0$ for exactly one positive frequency $\omega = \omega_n > 0$ and its negative $\omega = -\omega_n$ (there must be a zero for every n because (B.4) implies an upper bound in terms of the minimum of f on Σ , $E_n(\omega) < E_n(0) - \frac{1}{2}\omega^2 \min f$, which is negative for sufficiently large ω).

It is now straightforward to identify invariant and co-invariant states. Invariants are linear combinations of the positive and negative frequency modes solving the constraints,

$$\psi(T, X) = \frac{1}{\sqrt{2\pi}} \sum_n (\psi_n^+ e^{-iT\omega_n} \phi_n^{\omega_n}(X) + \psi_n^- e^{+iT\omega_n} \phi_n^{-\omega_n}(X)), \quad (\text{B.5})$$

with complex coefficients ψ_n^\pm . Co-invariants are given by smooth functions $f_n(\omega)$, with two functions $f \sim g$ identified as the same state if $f_n(\pm\omega_n) = g_n(\pm\omega_n)$ for all n .³⁸ The natural pairing between the two is

$$\langle\langle f|\psi\rangle\rangle = \sum_n (f_n^*(\omega_n)\psi_n^+ + f_n^*(-\omega_n)\psi_n^-). \quad (\text{B.6})$$

The decomposition into ‘Fourier modes’ $\hat{f}_n(\omega)$ also allows us to easily write the action of the rigging maps η (either group average or Klein-Gordon). For the group average, we have

$$\begin{aligned} \delta(H)f(T, X) &= \frac{1}{\sqrt{2\pi}} \sum_n \int d\omega \delta(E_n(\omega)) \hat{f}_n(\omega) e^{-iT\omega} \phi_n^\omega(X) \\ &= \frac{1}{\sqrt{2\pi}} \sum_n \frac{1}{|E'_n(\omega_n)|} (\hat{f}_n(\omega_n) e^{-iT\omega_n} \phi_n^{\omega_n}(X) + \hat{f}_n(-\omega_n) e^{+iT\omega_n} \phi_n^{-\omega_n}(X)). \end{aligned} \quad (\text{B.7})$$

So, we see that η projects the wavefunction onto the modes satisfying the constraint $E = 0$, with the positive factor $|E'_n(\omega_n)|^{-1}$ ((B.4) ensures that $E'_n(\omega_n)$ is nonzero). The resulting inner products on co-invariants and invariants give us

$$\langle\langle g|f\rangle\rangle := \langle g|\delta(H)|f\rangle = \sum_n \frac{\hat{f}_n(\omega_n)\hat{g}_n^*(\omega_n) + \hat{f}_n(-\omega_n)\hat{g}_n^*(-\omega_n)}{|E'_n(\omega)|}, \quad (\text{B.8})$$

$$\langle\langle \psi|\psi\rangle\rangle = \langle \psi|\kappa|\psi\rangle = \sum_n |E'_n(\omega)| (|\psi_n^+|^2 + |\psi_n^-|^2), \quad (\text{B.9})$$

which is manifestly positive definite.

We can compare this to the Klein-Gordon form, which corresponds to $\eta_{KG} = \frac{1}{2\pi i} \Delta$ given by the causal propagator $\Delta = \mathcal{G}_+ - \mathcal{G}_-$. For this, we can first write the retarded and advanced solutions $\mathcal{G}_\pm f$ to $H\psi = f$ as

$$\mathcal{G}_\pm f(T, X) = \frac{1}{\sqrt{2\pi}} \sum_n \int d\omega \frac{1}{E_n(\omega) \pm i\epsilon E'_n(\omega)} \hat{f}_n(\omega) e^{-iT\omega} \phi_n^\omega(X). \quad (\text{B.10})$$

These are manifestly solutions, and the $i\epsilon$ prescription ensures that the poles in ω are located in the upper or lower half-plane for retarded and advanced propagators respectively, yielding the appropriate boundary conditions. The difference between these gives a δ -function, but with a sign depending on the $i\epsilon$:

$$\begin{aligned} \Delta_n(\omega) &= \frac{1}{E_n(\omega) + i\epsilon E'_n(\omega)} - \frac{1}{E_n(\omega) - i\epsilon E'_n(\omega)} \\ &= -2\pi i \operatorname{sgn}(E'_n(\omega)) \delta(E_n(\omega)). \end{aligned} \quad (\text{B.11})$$

The resulting Klein-Gordon form on co-invariants is

$$\begin{aligned} \langle\langle g|f\rangle\rangle_{KG} &:= \frac{1}{2\pi i} \langle g|\Delta|f\rangle = - \sum_n \frac{\hat{f}_n(\omega_n)\hat{g}_n^*(\omega_n) + \hat{f}_n(-\omega_n)\hat{g}_n^*(-\omega_n)}{E'_n(\omega)} \\ &= \sum_n \frac{\hat{f}_n(\omega_n)\hat{g}_n^*(\omega_n) - \hat{f}_n(-\omega_n)\hat{g}_n^*(-\omega_n)}{|E'_n(\omega)|}, \end{aligned} \quad (\text{B.12})$$

³⁸We should also include technical conditions: $f_n(\omega)$ should be analytic of exponential type to ensure compactness in time (though a weaker condition suffices), and we should have growth conditions on ψ_n^\pm and f_n at large n . There is some freedom in the precise conditions, though they are replaced with weaker and more rigid requirements once we have an inner product and pass to the Hilbert space completion.

where in the last line we use the fact that E_n is decreasing for positive ω . This is almost identical to co-invariants, up to a crucial sign. It gives us a non-degenerate but indefinite Hermitian form: a positive norm for the positive-frequency modes, but negative frequency modes have negative norm.

We can give the sign appearing in the Klein-Gordon form a classical interpretation, by noting it comes from the sign of $E'_n(\omega)$ in the $i\epsilon$ prescription. This derivative of energy with respect to ω is classically equal to the derivative of target space time T with respect to worldline proper time τ , $\frac{dT}{d\tau} = \{T, H\} = -\frac{\partial H}{\partial \omega}$ (using the fact that $-\omega$ is the momentum conjugate to T). Positive frequency then corresponds to $\frac{dT}{d\tau} > 0$ and negative frequency $\frac{dT}{d\tau} < 0$. Later we will see how to generalise this idea, interpreting the sign of the norm as the direction in which the solution crosses the slice Σ in phase space, which will allow us to write more general expressions for both group average and Klein-Gordon type inner products on invariants.

C More on BRST calculations

Our discussion of inner products in the language of BFV or non-minimal BRST in section 6 relied heavily on exponentiating the anti-commutator of the BRST charge with some gauge fixing fermion Ψ . Here we will briefly elucidate some of the finer details which led to our results in the hope that it may aid with future calculations done by other authors. We first present an identity which is often useful for exponentiating $[Q, \Psi]_+$ in relatively simple cases.

Let ϵ be an operator which squares to zero $\epsilon^2 = 0$ and commutes with operators X, Y (for our applications, ϵ is a product of ghosts such as $\hat{c}\hat{b}$). Then

$$e^{X+\epsilon Y} = e^X + \epsilon \int_0^1 ds e^{(1-s)X} Y e^{sX}. \quad (\text{C.1})$$

We can think of this as taking the linear term in ϵ and averaging over all the possible orderings between e^X and Y .

To prove this, write the exponentials on both sides as power series and compare terms. For the left hand side we use

$$(X + \epsilon Y)^n = X^n + \epsilon \sum_{k=1}^n X^{k-1} Y X^{n-k}, \quad (\text{C.2})$$

so $\epsilon X^p Y X^q$ appears in the n th term of the exponential with $n = p+q+1$, giving a coefficient $\frac{1}{(p+q+1)!}$. For the right the same term appears with coefficient

$$\int_0^1 ds \frac{(1-s)^p}{p!} \frac{s^q}{q!}, \quad (\text{C.3})$$

which evaluates to the same. Use of this particular identity is useful in cases where $[Q, \Psi]_+$ has no terms involving conjugate pairs of ghosts and ghost momenta (i.e. $\hat{c}\hat{b}$ or $\hat{c}\hat{b}$).

When computing the inner products between invariants, we take $\Psi = -i\hat{b}\chi$, where χ is some ‘gauge-fixing’ operator on the matter Hilbert space. We get $[Q, \Psi]_+ = -i\Pi\chi -$

$i[H, \chi]c\hat{b}$. Using the identity proven above we find

$$e^{[Q, \chi]_+} = e^{-i\Pi\chi} - ic\hat{b} \int ds e^{-i(1-s)\Pi\chi} [H, \chi] e^{-is\Pi\chi}. \quad (\text{C.4})$$

In the main text we also make liberal use of the fact that,

$$[H, e^{-i\Pi\chi}] = -i\Pi \int ds e^{-i(1-s)\Pi\chi} [H, \chi] e^{-is\Pi\chi}. \quad (\text{C.5})$$

This can similarly be proven by Taylor expanding the exponential on either side. On the left hand side we get,

$$\begin{aligned} [H, e^{-i\Pi\chi}] &= \sum_{n=0}^{\infty} \frac{(-i\Pi)^n}{n!} [H, \chi^n] = \sum_{p,q=0}^{\infty} \frac{(-i\Pi)^{p+q+1}}{(p+q+1)!} \chi^p [H, \chi] \chi^q \\ &= -i\Pi \int_0^1 ds e^{-i(1-s)\Pi\chi} [H, \chi] e^{-is\Pi\chi}, \end{aligned} \quad (\text{C.6})$$

Which is precisely the right hand side.

We can additionally generalize the identity (C.1) slightly to the case where the operator in the exponential (the anti-commutator of the BRST charge with the gauge fixing fermion) does not just have a nilpotent term proportional to $c\hat{b}$ or $\hat{c}b$ but also has a term like cb or $\hat{c}\hat{b}$ (e.g., for the discussion above (6.10)). The cases we'd like to consider are when the operator we're exponentiating takes one of the two following forms,

$$\begin{aligned} \mathcal{O} &= bcX + cbY + \hat{c}bZ \\ &\text{or} \\ \tilde{\mathcal{O}} &= \hat{b}\hat{c}\tilde{X} + \hat{c}\hat{b}\tilde{Y} + \hat{c}\hat{b}\tilde{Z}, \end{aligned} \quad (\text{C.7})$$

Where X, Y, Z and their tilded variants are all taken to be Hermitian operators on the part of the factor of the Hilbert space without the ghosts. These operators arise naturally in contexts where we want to compute inner products between invariants using a gauge fixing fermion which does not act as the identity operator on the lapse degrees of freedom, or when we want to compute inner products between coinvariants which does not act as the identity operator on unconstrained Hilbert space.

Note that the presence of non-nilpotent combinations of ghosts means that we cannot straightforwardly apply equation (C.1). However, for the purpose of the inner product calculations for which this problem is relevant the important quantities will be the exponentiation of these operators evaluated in the $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ states (corresponding to coinvariants and invariants) respectively. As such we will specialize to these cases. As before, we will write the matrix exponentials in a series form

$$\begin{aligned} \langle\uparrow\downarrow|e^{\mathcal{O}}|\uparrow\downarrow\rangle &= \sum_n \frac{1}{n!} \langle\uparrow\downarrow|(bcX + cbY + \hat{c}bZ)^n|\uparrow\downarrow\rangle \\ \langle\downarrow\uparrow|e^{\tilde{\mathcal{O}}}\downarrow\uparrow\rangle &= \sum_n \frac{1}{n!} \langle\downarrow\uparrow|(\hat{b}\hat{c}\tilde{X} + \hat{c}\hat{b}\tilde{Y} + \hat{c}\hat{b}\tilde{Z})^n|\downarrow\uparrow\rangle. \end{aligned} \quad (\text{C.8})$$

Now we note that since $cb|\uparrow\downarrow\rangle = |\uparrow\downarrow\rangle$, $\hat{c}b|\uparrow\downarrow\rangle = bc|\downarrow\uparrow\rangle$, and $\hat{c}b|\downarrow\uparrow\rangle = bc|\uparrow\downarrow\rangle = cb|\downarrow\uparrow\rangle = 0$ we find that

$$\langle\uparrow\downarrow|(bcX + cbY + \hat{c}bZ)^n|\uparrow\downarrow\rangle = \langle\uparrow\downarrow|\downarrow\uparrow\rangle \times \sum_{i+j+1=n} Y^i Z X^j, \quad (\text{C.9})$$

and by similar reasoning, we find that

$$\langle\downarrow\uparrow|(\hat{b}\hat{c}\tilde{X} + \hat{c}\hat{b}\tilde{Y} + \hat{c}\hat{b}\tilde{Z})^n|\downarrow\uparrow\rangle = \langle\downarrow\uparrow|\uparrow\downarrow\rangle \times \sum_{i+j+1=n} \tilde{Y}^i \tilde{Z} \tilde{X}^j. \quad (\text{C.10})$$

Using these results, we arrive at the conclusion that

$$\begin{aligned} \langle\uparrow\downarrow|e^{\mathcal{O}}|\uparrow\downarrow\rangle &= \langle\uparrow\downarrow|\downarrow\uparrow\rangle \times \sum_{i,j} \frac{Y^i Z X^j}{(i+j+1)!} = \langle\uparrow\downarrow|\downarrow\uparrow\rangle \times \int_0^1 ds e^{(1-s)Y} Z e^{sX} \\ \langle\downarrow\uparrow|e^{\tilde{\mathcal{O}}}\downarrow\uparrow\rangle &= \langle\downarrow\uparrow|\uparrow\downarrow\rangle \times \sum_{i,j} \frac{\tilde{Y}^i \tilde{Z} \tilde{X}^j}{(i+j+1)!} = \langle\downarrow\uparrow|\uparrow\downarrow\rangle \times \int_0^1 ds e^{(1-s)\tilde{Y}} \tilde{Z} e^{s\tilde{X}}. \end{aligned} \quad (\text{C.11})$$

D Semi-classical calculations

D.1 Perturbative solution of WDW and inner product

In the u variable, we have

$$\psi''(u) - u\psi(u) = \sum_{n=2}^{\infty} \frac{V^{(n)}(q_0)}{n!V'(q_0)} \frac{u^n}{\lambda^{n-1}} \psi(u). \quad (\text{D.1})$$

Go to momentum space by writing an Ansatz

$$\psi(u) = \int \frac{dv}{2\pi} e^{ivu+iv^3/3} f(v), \quad (\text{D.2})$$

which gives us

$$f'(v) = \sum_{n=2}^{\infty} \frac{V^{(n)}(q_0)}{n!V'(q_0)} \frac{i^{n+1}}{\lambda^{n-1}} e^{-iv^3/3} \partial_v^n \left(e^{iv^3/3} f(v) \right). \quad (\text{D.3})$$

Now we simply expand $f(v) = \sum_{n=0}^{\infty} \lambda^{1-n} f_n(v)$ with $f_0(v) = 1$, and to n th order in λ we only get finitely many terms:

$$f'_n(v) = \sum_{k=2}^{n+1} \frac{V^{(k)}(q_0)}{k!V'(q_0)} \frac{i^{k+1}}{\lambda^{k-1}} e^{-iv^3/3} \partial_v^k \left(e^{iv^3/3} f_{n-k+1}(v) \right). \quad (\text{D.4})$$

At each order, the right hand side is a polynomial which we can easily integrate with $f_n(0) = 0$ for $n \geq 1$ to find f_n (a polynomial in v of degree $5n$). Furthermore, only coefficients of q^{5n-3k} for integer k are nonzero.

Now let's compute the inner product. We can expand it as

$$\frac{V(q_1) - V(q_2)}{q_1 - q_2} = \sum_{k_1, k_2=0}^{\infty} \frac{V^{(k_1+k_2+1)}(q_0)}{(k_1 + k_2 + 1)!} q_1^{k_1} q_2^{k_2}, \quad (\text{D.5})$$

and calculate terms from $\int q^k \psi(q) dq = \frac{i^k}{\lambda^{k+1}} \partial_v^k \left(e^{iv^3/3} f(v) \right) \Big|_{v=0}$. Acting on f_n , only derivatives where $k-n$ is a multiple of 3 give a nonzero result, so this has an expansion in powers of λ^{-3} , or \hbar^2 . We find

$$\|\psi\|^2 = (-V'(q_0)) |\tilde{\psi}(p=0)|^2 \left(1 + \frac{3V''(q_0)^3 + V^{(4)}(q_0)V'(q_0)^2 - 4V^{(3)}(q_0)V'(q_0)V'''(q_0)}{6V'(q_0)^3} \lambda^{-3} + \dots \right). \quad (\text{D.6})$$

D.2 Examples

Our first example is an exponential potential,

$$V(q) = \frac{\omega^2}{2}(1 - e^{\mu q}). \quad (\text{D.7})$$

The normalisable solution to the Wheeler-DeWitt equation is

$$\psi(q) = K \frac{2i\omega}{\hbar\mu} \left(\frac{2\omega}{\hbar\mu} e^{\frac{\mu q}{2}} \right), \quad (\text{D.8})$$

which we have not normalised. The $q \rightarrow -\infty$ asymptotics are

$$\psi(q) \sim \frac{1}{2} e^{\frac{2i\omega}{\hbar\mu} \log \frac{\omega}{\hbar\mu}} \Gamma\left(\frac{-2i\omega}{\hbar\mu}\right) e^{\frac{i\omega q}{\hbar}} + \text{c.c.} = \sqrt{\frac{\pi\mu\hbar}{2\omega \sinh\left(\frac{2\pi\omega}{\hbar\mu}\right)}} \cos\left(\frac{\omega q}{\hbar} + \theta\right), \quad (\text{D.9})$$

so from this we can read off the group-averaged norm of the state (using section 4.5) to be

$$\|\psi\|^2 = \frac{\pi\mu\hbar^2}{8 \sinh\left(\frac{2\pi\omega}{\hbar\mu}\right)}. \quad (\text{D.10})$$

The momentum space wavefunction is given by the Fourier transform

$$\psi(p) = \int \frac{dq}{\sqrt{2\pi\hbar}} \psi(q) e^{-ipq/\hbar} = \frac{1}{2\mu\sqrt{2\pi\hbar}} \left(\frac{\mu\hbar}{\omega}\right)^{-\frac{2ip}{\mu\hbar}} \Gamma\left(\frac{p-\omega}{i\mu\hbar}\right) \Gamma\left(\frac{p+\omega}{i\mu\hbar}\right), \quad (\text{D.11})$$

so in particular we have

$$\psi(p=0) = \sqrt{\frac{\pi\hbar}{2}} \frac{1}{2\omega \sinh\left(\frac{\pi\omega}{\hbar\mu}\right)}. \quad (\text{D.12})$$

Now, if we normalise this to get an in-state with unit norm we have

$$\psi_{\text{in}}(p=0) = \sqrt{\frac{2\pi}{\omega}} \frac{1}{2\omega \sinh\left(\frac{\pi\omega}{\hbar\mu}\right)} \frac{e^{\frac{2i\omega}{\hbar\mu} \log \frac{\omega}{\hbar\mu}}}{\Gamma\left(\frac{2i\omega}{\hbar\mu}\right)} \sim \frac{e^{\frac{i}{\hbar} \frac{2\omega}{\mu} (1-\log 2)}}{\sqrt{-i\hbar \frac{1}{2} \omega^2 \mu}} (1 + O(\hbar)), \quad (\text{D.13})$$

in precise agreement with (7.11). Note that as well as perturbative corrections (from the Γ function), expanding the sinh in the denominator gives us non-perturbative corrections suppressed by factors of $e^{-\frac{2\pi\omega}{\hbar\mu}}$. These can be accounted for by complex saddle points with paths ending at $q = \frac{2\pi i n}{\mu}$ where $V'(q) = 0$.

Now comparing the exact norm with the leading order result from the $p = 0$ gauge-fixed norm, we find

$$\frac{\hbar(-V'(0))|\psi(p=0)|^2}{\|\psi\|^2} = \frac{1}{\tanh\left(\frac{\pi\omega}{\hbar\mu}\right)}, \quad (\text{D.14})$$

which is one plus non-perturbatively small corrections. So for this potential, the leading order expression for the $p = 0$ gauge-fixed inner product is correct to all orders in perturbation theory, but not exact.

Another example is useful to show that the absence of perturbative corrections is special to this example. Take

$$V(q) = \frac{\omega^2}{2} \left(1 - \frac{a^2}{q^2} \right), \quad (\text{D.15})$$

where we restrict to $q < 0$ (and demand $\psi(q) \rightarrow 0$ at $q \rightarrow 0$ as the unique boundary conditions to get a self-adjoint Hamiltonian). The solution to WDW is

$$\psi(q) = \frac{1}{\hbar} \sqrt{-2\pi q} J_n \left(-\frac{\omega}{\hbar} q \right), \quad (\text{D.16})$$

where $n^2 = \frac{a^2 \omega^2}{\hbar^2} + \frac{1}{4}$, and we have normalised so that $\|\psi\| = 1$ using group-averaging (by looking at the $q \rightarrow -\infty$ tails). The zero momentum component (which we obtained from integrating $(-q)^{-k} \psi(q)$ for $k > 0$ and taking $k \rightarrow 0$) is

$$\psi(p=0) = \sqrt{\frac{2}{\omega^3} \frac{\Gamma\left(\frac{n}{2} + \frac{3}{4}\right)}{\Gamma\left(\frac{n}{2} + \frac{1}{4}\right)}} \sim \sqrt{\frac{a}{\hbar \omega^2} \left(1 + \frac{\hbar^2}{8a^2 \omega^2} + \dots \right)}. \quad (\text{D.17})$$

This gives us

$$\|\psi\|^2 = (-\hbar V'(a)) |\psi(p=0)|^2 \left(1 - \frac{\hbar^2}{4a^2 \omega^2} + \dots \right), \quad (\text{D.18})$$

with the expansion matching the perturbative corrections computed in the main text.

E String worldsheet analogy

While many of the ingredients we are discussing may seem unconventional, any reader familiar with the basics of string theory will already have encountered a version of all of them in the quantisation of the string worldsheet (see [38], for example). Since this analogy may be a helpful guide in a more familiar context, we sketch the main ideas here.

Specifically, consider the quantisation of a closed bosonic string in conformal gauge (similar comments apply to open strings and superstrings). We are left with the Hilbert space of a worldsheet CFT (e.g., 26 free bosons), but must still impose the constraints from the residual conformal gauge transformations generated by L_n, \bar{L}_n . The inner product on the space of physical string states is not usually discussed (since in this context the more relevant interpretation is in terms of the target space, rather than physics from the perspective of the worldsheet), but our considerations are nonetheless closely related to certain familiar amplitudes. Using ‘old covariant quantisation’ to identify the physical Hilbert space, for non-zero n we impose ‘invariance’ $L_n|\psi\rangle = 0$ on the positive-frequency modes $n > 0$, and ‘co-invariance’ equivalence relations on the negative-frequency modes $n < 0$.³⁹ We will not say more about these constraints, concentrating on the remaining zero modes $H = L_0 + \bar{L}_0 - 2$ and $P = L_0 - \bar{L}_0$. The momentum P generates the compact $U(1)$ group of translations of the string, so we can simply project onto the $P = 0$ subspace (or average over the finite-volume group) and this does not lead to any subtleties with

³⁹This arises in BRST quantisation by choosing a state for the corresponding ghosts which is annihilated by b_n and c_n for $n < 0$.

constructing a gauge-invariant inner product on the space of physical states. The only tricky part concerns the Hamiltonian constraint H (and this works essentially the same as the single constraint H in the one-dimensional models that this paper concentrates on).

The first (and most standard) way to define physical states imposes ‘invariance’ for the zero modes also, $H|\psi\rangle = P|\psi\rangle = 0$. This gives rise to the usual string states corresponding to marginal ($h = \bar{h} = 1$) vertex operators of the CFT. But in fact we also encounter states of the string which we might try to interpret as co-invariants with respect to H : these are boundary states, representing a string emitted from a D-brane. These states are annihilated by P , but we do not require that they are annihilated by H (and the same applies for all modes $L_n + \bar{L}_n$ of the Hamiltonian constraints).⁴⁰

Now, even for non-interacting strings (at zero string coupling) we would expect to be able to construct an inner product (or at least a gauge-invariant amplitude) from a pair of such physical states. The simplest case is when we take one of each: a vertex operator state $|\mathcal{V}\rangle$ and boundary state $|\mathcal{B}\rangle$ have a natural finite gauge-invariant pairing given by the disk one-point function. This amplitude is particularly simple since it has no moduli to integrate over, and no residual gauge symmetry. It is analogous to the canonical pairing $\langle\langle \mathcal{B}|\mathcal{V} \rangle\rangle$ between invariants and co-invariants, and the simplicity of the amplitude reflects that this pairing is natural, requiring no further choices to be made.

The pairing of two vertex operators \mathcal{V}_1 and \mathcal{V}_2 gives us the sphere two-point amplitude, but this is more complicated because there is an infinite-volume residual symmetry generated by H (as well as the finite-volume rotation generated by P). This comes about because vertex operator states are not normalisable due to the continuous spectrum of H . For a string in D non-compact spacetime dimensions ($D = 26$ in the simplest case), the naïve inner product is proportional to $\delta^{(D)}(k_\mu - k'_\mu)$, a delta-function in the full spacetime. The physically sensible result includes only a delta-function $\delta^{(D-1)}(\vec{k} - \vec{k}')$ in the spatial directions, with the last equality automatically following from the mass-shell condition $p^2 + m^2 = 0$. To get a sensible finite result requires some gauge-fixing condition (such as $X^0 = 0$ at a chosen point on the worldsheet) [39] (which is associated with a c -ghost insertion in the BRST formalism). This is completely analogous to the Klein-Gordon-like κ map.

Finally, for two boundary states $\mathcal{B}_1, \mathcal{B}_2$ the obvious pairing requires the cylinder amplitude (or annulus). In this case there is a modulus (the length of the cylinder for fixed radius), a single-parameter family of gauge-inequivalent worldsheet conformal structures to integrate over (associated with a b -ghost insertion). We get an amplitude of the form $\int dt \langle \mathcal{B}_2 | e^{-iHt} | \mathcal{B}_1 \rangle$, which should be compared to the group-averaging integral in (1.6). There is an important difference for the usual string amplitude: we integrate over imaginary t (Euclidean worldsheets), and only over a half-line rather than all real t . This is because the string amplitude has a target space interpretation, and is not intended as an inner product

⁴⁰Keeping track of the ghosts, the vertex operator states are annihilated by b_0, \bar{b}_0 (Seigel gauge), as appropriate for invariants. In the path integral, fixing the coordinate location of vertex operators leads to a ghost operator insertion $c\tilde{c}$, which under the state operator correspondence leads to the correct state. Boundary states are annihilated by $b_n - \tilde{b}_n$, which are the b -ghosts associated with momentum constraints, but by $c_n + \tilde{c}_n$ corresponding to Hamiltonian constraints.

Gravitational Hilbert space	String worldsheet Hilbert space
Invariant states $ \psi\rangle$	Vertex operators
Co-invariant states $ \psi\rangle\rangle$	Boundary states (D-branes)
Pairing $\llbracket\psi_2 \psi_1\rangle$ between invariants	Sphere two-point amplitude
Pairing $\llbracket\psi_2 \psi_1\rangle\rangle$ between invariants and co-invariants	Disk one-point amplitude
Pairing $\langle\langle\psi_2 \psi_1\rangle\rangle$ between co-invariants	Cylinder/annulus amplitude

Table 1. Analogies between the gravitational Hilbert space and that of the string worldsheet.

of worldsheet states (see comments in [27]).

F Maxwell theory

To illustrate some of the ideas in a more familiar setting, we here discuss time evolution and inner products in the BFV formalism for electrodynamics.

In terms of the (Lorentz non-covariant) electric scalar potential ϕ and magnetic vector potential \vec{A} , the Lagrangian of the theory is

$$L = \frac{1}{2} \int d^3x (E^2 - B^2), \quad (F.1)$$

$$\vec{B} = \nabla \times \vec{A}, \quad \vec{E} = -\frac{\partial \vec{A}}{\partial t} - \nabla \phi.$$

From this, the electric field \vec{E} is minus the canonical momentum for \vec{A} (and we will not introduce separate notation for this), while the momentum Π conjugate to ϕ vanishes: we have the primary constraint $\Pi = 0$. The canonical construction of the Hamiltonian gives (after an integration by parts)

$$H = \int d^3x \left[\frac{1}{2} (E^2 + B^2) - \phi \nabla \cdot E \right] \quad (\vec{B} = \nabla \times \vec{A}). \quad (F.2)$$

Demanding that this commutes with Π , we get the secondary constraint $\nabla \cdot E = 0$ (Gauss's law), and ϕ acts as the Lagrange multiplier imposing this constraint.

Since the two constraints $\Pi, \nabla \cdot E$ commute, the BRST charge is simply

$$Q = \int d^3x [c \nabla \cdot E + \hat{c} \Pi]. \quad (F.3)$$

This BRST charge does not commute with H as given above, signalling the need for additional ghost terms. But Q does commute with the first term, and so we can write the most general Hamiltonian as $H_\Psi = \frac{1}{2} \int (E^2 + B^2) + [Q, \Psi]_+$ for some choice of gauge-fixing fermion Ψ . A nice family of choices are the Gaussian-averaged Lorenz gauges (parameterised by a real number ξ):

$$\Psi = - \int d^3x (b\phi + \hat{b} \nabla \cdot A + \frac{1}{2} \xi \hat{b} \Pi). \quad (F.4)$$

This gives us a Hamiltonian $H_\Psi = H_\xi + H_{\text{ghost}}$, with

$$\begin{aligned} H_\xi &= \int d^3x \left[\frac{1}{2}(E^2 + B^2) - \phi \nabla \cdot E - \Pi \nabla \cdot A - \frac{1}{2}\xi \Pi^2 \right] \\ H_{\text{ghost}} &= i \int d^3x (\hat{c}b + c\nabla^2 \hat{b}). \end{aligned} \quad (\text{F.5})$$

The correspondence to Lorenz gauge is visible from the Hamiltonian equation of motion for ϕ (obtained from varying Π), which is $\dot{\phi} = \frac{\delta H}{\delta \Pi} = -\nabla \cdot A - \xi \Pi$, which we can write as $\xi \Pi = -\partial_\mu A^\mu$. For $\xi = 0$ this is the Lorenz gauge condition $\partial_\mu A^\mu = 0$ directly, while for $\xi \neq 0$ we get the Lorenz gauge condition after imposing the constraint $\Pi = 0$. More generally, the bosonic part H_ξ can also be obtained by adding a term $-\frac{1}{2\xi}(\partial_\mu A^\mu)^2$ to the original Lagrangian. Landau gauge is $\xi = 0$, and Feynman gauge is $\xi = 1$.

F.1 Wavefunctional representation

Let's understand the action of the time-evolution operator $e^{-iH_\xi t}$ in this gauge (since H_{ghost} is independent of the original fields we'll concentrate on the bosonic part for now, coming back to the ghost contribution later). This will be slightly easier in momentum space, defining $\phi(x) = \int \frac{d^3x}{(2\pi)^{3/2}} e^{ik \cdot x} \phi(k)$ (with $\phi(k)^\dagger = \phi(-k)$) and similarly for other fields, to get

$$H_\xi = \int d^3k \left[\frac{1}{2}|E(k)|^2 + \frac{1}{2}(k^2|A(k)|^2 - |k \cdot A(k)|^2) - \frac{1}{2}\xi|\Pi(k)|^2 - i\phi(-k)k \cdot E(k) - i\Pi(-k)k \cdot A(k) \right]. \quad (\text{F.6})$$

Now we can separate the physical transverse modes of the gauge field and electric field from the nonphysical longitudinal components: $\vec{A}(k) = \vec{A}_\perp(k) + \frac{\vec{k}}{|\vec{k}|}A_\parallel(k)$ where $A_\parallel(k) = \frac{1}{|\vec{k}|}\vec{k} \cdot \vec{A}(k)$ is the component of \vec{A} in the direction \vec{k} , and similarly for E_\parallel . Note that $A_\parallel(k)^\dagger = -A_\parallel(-k)$. The B^2 'potential' term for \vec{A} projects onto the physical transverse part, so we can write

$$\begin{aligned} H_\xi &= H_\perp + H_{\text{gauge}}, \\ H_\perp &= \int d^3k \left[\frac{1}{2}|E_\perp(k)|^2 + \frac{1}{2}k^2|A_\perp(k)|^2 \right], \\ H_{\text{gauge}} &= \int d^3k \left[\frac{1}{2}|E_\parallel(k)|^2 - \frac{1}{2}\xi|\Pi(k)|^2 - i|k|\phi(-k)E_\parallel(k) - i|k|\Pi(-k)A_\parallel(k) \right]. \end{aligned} \quad (\text{F.7})$$

The transverse piece is simply the free Hamiltonian for the physical components of the fields (with massless relativistic dispersion relation $\omega^2 = k^2$) and is unaffected by the constraints, so we can concentrate on the (commuting) 'gauge' piece, acting on the longitudinal modes and ϕ, Π .

Now, since different values of momentum k are all independent (except that modes with $-k$ and k are related by complex conjugation) we can treat them all separately, treating each \vec{k} (together with $-\vec{k}$) as an ordinary quantum mechanics with two complex coordinates A_\parallel and ϕ , with conjugate momenta $-E_\parallel$ and Π . To recover the full results we need only restore k -dependence and take appropriate integrals over k in the end. With this we can write the momentum $\pm\vec{k}$ contribution to H_{gauge} as

$$H_{\text{gauge}}(k) = |E_\parallel|^2 - \xi|\Pi|^2 - i|k|(\bar{\phi}E_\parallel - \phi\bar{E}_\parallel + \bar{\Pi}A_\parallel - \Pi\bar{A}_\parallel). \quad (\text{F.8})$$

For these complex coordinates the nontrivial commutation relations are $[\phi, \bar{\Pi}] = [\bar{\phi}, \Pi] = i$, and $[A_{\parallel}, \bar{E}_{\parallel}] = [\bar{A}_{\parallel}, E_{\parallel}] = -i$.

Since H_{gauge} is linear in A_{\parallel} and ϕ , it is simplest to solve the Schrödinger equation in the momentum space representation (with wavefunctions depending on E_{\parallel} , Π), so $\phi = i\partial_{\bar{\Pi}}$ and so forth. Then, H_{gauge} is a first-order differential operator,

$$H_{\text{gauge}} = |E_{\parallel}|^2 - \xi|\Pi|^2 + k \left(E_{\parallel} \frac{\partial}{\partial \Pi} - \bar{E}_{\parallel} \frac{\partial}{\partial \bar{\Pi}} + \Pi \frac{\partial}{\partial E_{\parallel}} - \bar{\Pi} \frac{\partial}{\partial \bar{E}_{\parallel}} \right) \quad (\text{F.9})$$

so the Schrödinger equation can then be solved by the method of characteristics. The Schrödinger evolution for time t from initial wavefunction ψ_0 then gives

$$\begin{aligned} \psi_t(E_{\parallel}, \Pi) &= e^{i\vartheta} \psi_0(E_{\parallel} \cos(kt) - i\Pi \sin(kt), \Pi \cos(kt) - iE_{\parallel} \sin(kt)), \\ \vartheta &= \frac{\xi-1}{2} t (|E_{\parallel}|^2 + |\Pi|^2) + \frac{\xi+1}{2} \frac{\sin(kt)}{k} (\cos(kt)(|\Pi|^2 - |E_{\parallel}|^2) + i \sin(kt)(\bar{E}_{\parallel}\Pi - E_{\parallel}\bar{\Pi})). \end{aligned} \quad (\text{F.10})$$

Now we have to choose how we want to represent our initial and final states: invariants, co-invariants, or a mixture of the two. For co-invariants, we choose arbitrary E_{\parallel} dependence, but impose invariance under the primary constraint $\Pi = 0$: so the initial wavefunction is $\psi(E_{\parallel})\delta(\Pi)$ (the notation is not supposed to imply that ψ is holomorphic, and the δ -function is two-dimensional, defined for a complex value). Evolving this gives us a state

$$\psi_t(E_{\parallel}, \Pi) = \exp \left[i \left(\frac{\xi-1}{2} t \sec^2(kt) - \frac{\xi+1}{2} \frac{\tan(kt)}{k} \right) |E_{\parallel}|^2 \right] \delta(\Pi \cos(kt) - iE_{\parallel} \sin(kt)) \psi(\sec(kt)E_{\parallel}) \quad (\text{F.11})$$

where we have assumed that $\cos(kt)$ is nonzero to write the answer in this form (otherwise the δ -function sets E_{\parallel} to zero and the remaining dependence must be written in terms of Π). Taking the inner product with a similar $\Pi = 0$ state means that we evaluate at the final wavefunction at $\Pi = 0$. This gives simply

$$\psi_t(E_{\parallel}(k), 0) = \left(\frac{\sin(kt)}{k} \right)^{-2} \delta(kE_{\parallel}) \psi(0), \quad (\text{F.12})$$

where the prefactor is a Jacobian coming from extracting a factor from the delta-function $\delta(-i \sin(kt)E_{\parallel})$. This factor should be cancelled by the contribution of the ghosts. Now kE_{\parallel} is simply a momentum space representation of the Gauss law constraint $\nabla \cdot E$, so our result gives us matrix elements of $\eta \propto \delta[\nabla \cdot E]$, the group averaged inner product projecting onto solutions of the constraints.

The Jacobian should be cancelled by a contribution from the ghosts. Following the same momentum space representation, we can write the momentum k ghost Hamiltonian as

$$H_{\text{ghost}}(k) = i(\hat{c}^\dagger b - k^2 c^\dagger \hat{b}) + i(\hat{c} b^\dagger - k^2 c \hat{b}^\dagger), \quad (\text{F.13})$$

with $[c, b^\dagger]_+ = [c^\dagger, b]_+ = 1$ and similarly for hatted ghosts. The two terms we have written commute, which is useful for calculations. For the co-invariant inner product, we want to calculate the matrix elements of $e^{-itH_{\text{ghost}}(k)}$ between states $|\uparrow\uparrow\downarrow\downarrow\rangle$ annihilated by $c, c^\dagger, \hat{b}, \hat{b}^\dagger$. To do this, simply use the Taylor expansion for the exponentials and note that the only

nonzero contribution comes from odd powers with alternating terms (meaning products like $(\hat{c}^\dagger b c^\dagger \hat{b})^{2n} \hat{c}^\dagger b$). These sum to give

$$\langle \uparrow\uparrow\downarrow\downarrow | e^{-itH_{\text{ghost}}(k)} | \uparrow\uparrow\downarrow\downarrow \rangle = \left(\frac{\sin(kt)}{k} \right)^2 \langle \uparrow\uparrow\downarrow\downarrow | \downarrow\downarrow\uparrow\uparrow \rangle, \quad (\text{F.14})$$

precisely cancelling the Jacobian factor from the bosonic part.

If we instead use invariant states, then we have initial and final wavefunctions $\delta(kE_{\parallel})\psi(\Pi)$ for arbitrary $\psi(\Pi)$. Evolving this gives

$$\psi_t(E_{\parallel}, \Pi) = \exp \left[i \left(\frac{\xi-1}{2} t \sec^2(kt) + \frac{\xi+1}{2} \frac{\tan(kt)}{k} \right) |\Pi|^2 \right] \delta(E_{\parallel} k \cos(kt) - i\Pi k \sin(kt)) \psi \left(\frac{\Pi}{\cos(kt)} \right). \quad (\text{F.15})$$

Taking the overlap with $\delta(kE_{\parallel})$, we get

$$\frac{1}{k} \psi_t(0, \Pi) = (k \sin(kt))^{-2} \delta(\Pi) \psi(0). \quad (\text{F.16})$$

Similarly to above, the ghost contribution is $(k \sin(kt))^2$, cancelling the Jacobian factor as expected.

F.2 Oscillator representation and Gupta-Bleuler

To explain the relation to the more familiar Lorenz gauge Gupta-Bleuler quantisation, we write the fields in terms of oscillators. Define annihilation operators φ_k and \vec{a}_k (which is vector-valued) and their conjugate creation operators as follows:

$$\begin{aligned} \phi(x) &= \int \frac{d^3k}{(2\pi)^{3/2}} \frac{e^{ik \cdot x}}{\sqrt{2|k|}} \varphi_k + \text{h.c.} \\ \vec{A}(x) &= \int \frac{d^3k}{(2\pi)^{3/2}} \frac{e^{ik \cdot x}}{\sqrt{2|k|}} \vec{a}_k + \text{h.c.} \\ \Pi(x) &= \frac{2i}{\xi+1} \int \frac{d^3k}{(2\pi)^{3/2}} e^{ik \cdot x} \sqrt{\frac{|k|}{2}} (\varphi_k - a_k^{\parallel}) + \text{h.c.} \\ \vec{E}(x) &= i \int \frac{d^3k}{(2\pi)^{3/2}} e^{ik \cdot x} \sqrt{\frac{|k|}{2}} \left(\vec{a}_k^{\perp} - \frac{2}{\xi+1} (\varphi_k - a_k^{\parallel}) \frac{\vec{k}}{|k|} \right) + \text{h.c.} \end{aligned} \quad (\text{F.17})$$

Here $a_k^{\parallel} := \frac{\vec{k} \cdot \vec{a}_k}{|k|}$ is the longitudinal component of \vec{a}_k , and $\vec{a}_k^{\perp} = \vec{a}_k - a_k^{\parallel} \frac{\vec{k}}{|k|}$ the transverse part. This decomposition is chosen so that the explicitly written pieces are ‘positive frequency’: Heisenberg evolution with H_{ξ} gives the time-dependence

$$\begin{aligned} \vec{a}_k(t) &= e^{-i|k|t} \left(\vec{a}_k - i \frac{\xi-1}{\xi+1} (\varphi_k - a_k^{\parallel}) \frac{\vec{k}}{|k|} t \right), \\ \varphi_k(t) &= e^{-i|k|t} \left(\varphi_k - i \frac{\xi-1}{\xi+1} (\varphi_k - a_k^{\parallel}) |k| t \right), \end{aligned} \quad (\text{F.18})$$

which is almost positive frequency oscillation (slightly spoiled by the terms linear in t). Here we see that Feynman gauge $\xi = 1$ simplifies things significantly, so we specialise to that case. We have non-trivial commutators

$$[\varphi_k, \varphi_{k'}^\dagger] = -\delta(k - k'), \quad [a_k^i, (a_{k'}^j)^\dagger] = \delta_{ij} \delta(k - k'), \quad (\text{F.19})$$

so that \bar{a}_k are ordinary annihilation operators, while the extra sign in the first relation means that (for the algebra at least) φ_k behaves as a creation operator. We can combine these back into a four-vector a_k^μ with $\varphi_k = a_k^0$, with covariant commutation relations $[a_k^\mu, (a_{k'}^\nu)^\dagger] = \eta^{\mu\nu} \delta(k - k')$. In terms of the oscillators, the Hamiltonian (up to a constant from operator ordering) is

$$H_{\xi=1} = \int d^3k |k| \left(\bar{a}_k^\dagger \cdot \bar{a}_k - \varphi_k^\dagger \varphi_k \right). \quad (\text{F.20})$$

Now, a quantisation of the oscillator commutation relations with a unitary representation on the Hilbert space must treat φ_k^\dagger as an annihilation operator, so there is a ‘vacuum’ state annihilated by a_k^i and φ_k^\dagger and other states are built by acting with $(a_k^i)^\dagger$ and φ_k . Then H_ξ would be unbounded from below (since φ_k lowers the energy). However, Gupta-Bleuler quantisation uses instead a non-unitary representation built on a Lorentz-invariant vacuum state $|\Omega\rangle$ annihilated by φ_k and a_k^i (and acting with $(a_k^i)^\dagger$ and φ_k^\dagger creates other states, always increasing energy by $|k|$). Before imposing constraints, we have a ‘Hilbert space’ with a non-degenerate but indefinite inner product (for example, a single photon state with timelike polarisation $\varphi_k^\dagger |\Omega\rangle$ has negative norm). The negative norm states are not immediately a problem, but they should be eliminated when we impose constraints and pass to the physical Hilbert space.

We can make a similar mode decomposition of the ghosts, with complex ghost modes b_k and c_k defined as follows:

$$\begin{aligned} c(x) &= \int \frac{d^3k}{(2\pi)^{3/2}} \frac{e^{ik \cdot x}}{\sqrt{2|k|}} c_k + \text{h.c.} \\ b(x) &= \int \frac{d^3k}{(2\pi)^{3/2}} e^{ik \cdot x} \sqrt{\frac{|k|}{2}} b_k + \text{h.c.} \\ \hat{c}(x) &= i \int \frac{d^3k}{(2\pi)^{3/2}} e^{ik \cdot x} \sqrt{\frac{|k|}{2}} c_k + \text{h.c.} \\ \hat{b}(x) &= i \int \frac{d^3k}{(2\pi)^{3/2}} \frac{e^{ik \cdot x}}{\sqrt{2|k|}} b_k + \text{h.c.} \end{aligned} \quad (\text{F.21})$$

These definitions are chosen so that the Heisenberg evolution of b_k and c_k simply produces a factor of $e^{-i|k|t}$, a positive-frequency oscillation. The only nontrivial anti-commutation relations following from these definitions are $[b_k^\dagger, c_{k'}]_+ = \delta(k - k')$ (and the conjugate $[b_k, c_{k'}^\dagger]_+ = \delta(k - k')$). The BRST charge and the ghost Hamiltonian are given in terms of oscillators by

$$Q = \int d^3k |k| \left(c_k^\dagger (\varphi_k - a_k^\parallel) + c_k (\varphi_k - a_k^\parallel)^\dagger \right), \quad (\text{F.22})$$

$$H_{\text{ghost}} = \int d^3k |k| \left(c_k^\dagger b_k + b_k^\dagger c_k \right). \quad (\text{F.23})$$

This means that b_k and c_k both lower the energy by $|k|$, and b_k^\dagger and c_k^\dagger both raise it.

To recover the standard Gupta-Bleuler quantisation we place the ghosts in their ground state, annihilated by all the positive-frequency modes b_k and c_k . To get a BRST-closed state, we then require that the matter part of the state is annihilated by $\varphi_k - a_k^\parallel$: this is the Gupta-Bleuler condition that the state is annihilated by the positive-frequency part of $\partial_\mu A^\mu$ (or equivalently, by the positive-frequency part of Π or of $\nabla \cdot E$). This roughly means that every unphysical timelike-polarized photon is accompanied by another unphysical longitudinal-polarized photon. The BRST-exact states of this form, given by Q acting on a state with a single b_k^\dagger ghost excitation, give the image of $(\varphi_k - a_k^\parallel)^\dagger$ in the matter sector. Imposing equivalence under addition of such a state, we learn that we can remove a pair of timelike and longitudinal photons and get the same physical state. The upshot is that the physical content of the BRST cohomology leaves only the physical transverse polarizations.

In this case, we consistently get an inner product in which the ghost vacuum (annihilated by b_k, c_k) has non-zero norm, and the matter oscillator vacuum has finite norm. Additionally, the physical state Gupta-Bleuler condition eliminates negative norm states and the quotient by BRST exact states leaves a sensible positive-definite inner product for the physical Hilbert space. So, there is no requirement to insert a non-trivial gauge transformation in order to define a physical inner product, as we did for one-dimensional gravity.

However, this would not be quite so simple if we changed things slightly to have a zero mode: for example, if we put the theory on a compact space. In that case, in the BRST charge gets a zero-mode contribution of the form $c_0 Q_m + \hat{c}_0 \Pi_0$, where Q_m is the total electric charge of matter coupled to our gauge field (which we haven't considered until now). The relevant zero-mode piece of the Hamiltonian H_ξ is $-Q_m \phi_0 - \frac{\xi}{2} \Pi^2 + i \hat{c}_0 b_0$. Now the situation is extremely similar to that considered in the text, where we constructed the group-averaged inner product. We can take our states to be co-invariant with respect to the zero-mode of the Gauss-law constraint, namely the vanishing of the total charge $Q_m = 0$; this means taking the ghost state annihilated by c_0 and \hat{b}_0 and a wavefunction independent of the zero mode of the electric potential ϕ_0 (annihilated by Π_0), but a state of arbitrary charge. The norm of such states take the indefinite form $0 \times \infty$ from the ghost and ϕ_0 sectors (for a non-compact gauge group). But any amount of time-evolution by H_ξ resolves this, with the zero-modes contributing a factor $t \int \frac{d\phi_0}{2\pi} e^{itQ_m \phi_0} = \text{sgn}(t) \delta(Q_m)$ (for any ξ). This projects onto the zero-charge component of the state (with a delta-function as appropriate for continuous charge under a non-compact gauge group), by averaging over the group of zero-mode gauge transformations (constant gauge parameter). For a compact gauge group and quantised charge we would change this range of integration; the BRST formalism does not directly address such global properties of the gauge group.

References

- [1] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013, [[1911.12333](#)].

- [2] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, *JHEP* **03** (2022) 205, [[1911.11977](#)].
- [3] P. Saad, S. H. Shenker and D. Stanford, *JT gravity as a matrix integral*, [1903.11115](#).
- [4] L. V. Iliesiu, S. Murthy and G. J. Turiaci, *Black hole microstate counting from the gravitational path integral*, [2209.13602](#).
- [5] J. Held and H. Maxfield, *The Hilbert space of de Sitter JT: a case study for canonical methods in quantum gravity*, [2410.14824](#).
- [6] P. A. M. Dirac, *Generalized Hamiltonian dynamics*, *Can. J. Math.* **2** (1950) 129–148.
- [7] B. S. DeWitt, *Quantum Theory of Gravity. 1. The Canonical Theory*, *Phys. Rev.* **160** (1967) 1113–1148.
- [8] V. Chandrasekaran, R. Longo, G. Penington and E. Witten, *An algebra of observables for de Sitter space*, *JHEP* **02** (2023) 082, [[2206.10780](#)].
- [9] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao and T. Thiemann, *Quantization of diffeomorphism invariant theories of connections with local degrees of freedom*, *J. Math. Phys.* **36** (1995) 6456–6493, [[gr-qc/9504018](#)].
- [10] D. Giulini and D. Marolf, *On the generality of refined algebraic quantization*, *Class. Quant. Grav.* **16** (1999) 2479–2488, [[gr-qc/9812024](#)].
- [11] D. Marolf, *Group averaging and refined algebraic quantization: Where are we now?*, in *9th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (MG 9)*, 7, 2000. [gr-qc/0011112](#).
- [12] R. M. Wald, *A Proposal for solving the ‘problem of time’ in canonical quantum gravity*, *Phys. Rev. D* **48** (1993) R2377–R2381, [[gr-qc/9305024](#)].
- [13] D. Marolf, *Quantum observables and recollapsing dynamics*, *Class. Quant. Grav.* **12** (1995) 1199–1220, [[gr-qc/9404053](#)].
- [14] A. Higuchi, *Quantum linearization instabilities of de Sitter space-time. 1*, *Class. Quant. Grav.* **8** (1991) 1961–1981.
- [15] E. Witten, *A note on the canonical formalism for gravity*, *Adv. Theor. Math. Phys.* **27** (2023) 311–380, [[2212.08270](#)].
- [16] J. J. Halliwell and J. B. Hartle, *Wave functions constructed from an invariant sum over histories satisfy constraints*, *Phys. Rev. D* **43** (1991) 1170–1194.
- [17] G. Araujo-Regado, R. Khan and A. C. Wall, *Cauchy slice holography: a new AdS/CFT dictionary*, *JHEP* **03** (2023) 026, [[2204.00591](#)].
- [18] R. Marnelius and M. Ogren, *Symmetric inner products for physical states in BRST quantization*, *Nucl. Phys. B* **351** (1991) 474–490.
- [19] I. Batalin and R. Marnelius, *Solving general gauge theories on inner product spaces*, *Nucl. Phys. B* **442** (1995) 669–696, [[hep-th/9501004](#)].
- [20] O. Y. Shvedov, *On correspondence of BRST-BFV, Dirac and refined algebraic quantizations of constrained systems*, *Annals Phys.* **302** (2002) 2–21, [[hep-th/0111270](#)].
- [21] E. S. Fradkin and G. A. Vilkovisky, *QUANTIZATION OF RELATIVISTIC SYSTEMS WITH CONSTRAINTS*, *Phys. Lett. B* **55** (1975) 224–226.

- [22] I. A. Batalin and G. A. Vilkovisky, *Relativistic S Matrix of Dynamical Systems with Boson and Fermion Constraints*, *Phys. Lett. B* **69** (1977) 309–312.
- [23] M. Henneaux, *Hamiltonian Form of the Path Integral for Theories with a Gauge Freedom*, *Phys. Rept.* **126** (1985) 1–66.
- [24] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems*. Princeton University Press, 8, 1994.
- [25] J. J. Halliwell, *Derivation of the Wheeler-De Witt Equation from a Path Integral for Minisuperspace Models*, *Phys. Rev. D* **38** (1988) 2468.
- [26] I. Moss, *Quantum Cosmology and the Selfobserving Universe*, *Ann. Inst. H. Poincaré Phys. Theor.* **49** (1988) 341–349.
- [27] E. Casali, D. Marolf, H. Maxfield and M. Rangamani, *Baby universes and worldline field theories*, *Class. Quant. Grav.* **39** (2022) 134004, [[2101.12221](#)].
- [28] D. Marolf, *Refined algebraic quantization: Systems with a single constraint*, [gr-qc/9508015](#).
- [29] C. Bär, *Linear wave equations on lorentzian manifolds*, [1006.2354](#).
- [30] D. Harlow and D. Jafferis, *The Factorization Problem in Jackiw-Teitelboim Gravity*, *JHEP* **02** (2020) 177, [[1804.01081](#)].
- [31] L. V. Iliesiu, A. Levine, H. W. Lin, H. Maxfield and M. Mezei, *On the non-perturbative bulk Hilbert space of JT gravity*, *JHEP* **10** (2024) 220, [[2403.08696](#)].
- [32] R. P. Woodard, *Enforcing the Wheeler-de Witt Constraint the Easy Way*, *Class. Quant. Grav.* **10** (1993) 483–496.
- [33] D. Giulini and D. Marolf, *A Uniqueness theorem for constraint quantization*, *Class. Quant. Grav.* **16** (1999) 2489–2505, [[gr-qc/9902045](#)].
- [34] D. Marolf, *The Spectral analysis inner product for quantum gravity*, in *7th Marcel Grossmann Meeting on General Relativity (MG 7)*, pp. 851–853, 9, 1994. [gr-qc/9409036](#).
- [35] S. Hollands and R. M. Wald, *Quantum fields in curved spacetime*, *Phys. Rept.* **574** (2015) 1–35, [[1401.2026](#)].
- [36] R. M. Wald, *Quantum Field Theory in Curved Space-Time and Black Hole Thermodynamics*. Chicago Lectures in Physics. University of Chicago Press, Chicago, IL, 1995.
- [37] A. G. Cohen, G. W. Moore, P. C. Nelson and J. Polchinski, *An Off-Shell Propagator for String Theory*, *Nucl. Phys. B* **267** (1986) 143–157.
- [38] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 12, 2007, [10.1017/CBO9780511816079](#).
- [39] H. Erbin, J. Maldacena and D. Skliros, *Two-Point String Amplitudes*, *JHEP* **07** (2019) 139, [[1906.06051](#)].
- [40] D. Marolf, *Path integrals and instantons in quantum gravity: Minisuperspace models*, *Phys. Rev. D* **53** (1996) 6979–6990, [[gr-qc/9602019](#)].
- [41] J. B. Hartle and D. Marolf, *Comparing formulations of generalized quantum mechanics for reparametrization - invariant systems*, *Phys. Rev. D* **56** (1997) 6247–6257, [[gr-qc/9703021](#)].

- [42] A. Higuchi and R. M. Wald, *Applications of a new proposal for solving the 'problem of time' to some simple quantum cosmological models*, *Phys. Rev. D* **51** (1995) 544–561, [[gr-qc/9407038](#)].
- [43] B. S. DeWitt, *The Quantization of geometry*, .
- [44] S. B. Giddings, D. Marolf and J. B. Hartle, *Observables in effective gravity*, *Phys. Rev. D* **74** (2006) 064018, [[hep-th/0512200](#)].
- [45] V. N. Gribov, *Quantization of Nonabelian Gauge Theories*, *Nucl. Phys. B* **139** (1978) 1.
- [46] H. Arisue, T. Fujiwara, T. Inoue and K. Ogawa, *GENERALIZED SCHRODINGER REPRESENTATIONS AND ITS APPLICATION TO GAUGE THEORIES.*, in *1981 INS Symposium on Quark and Lepton Physics*, 1981.
- [47] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner and G. Vattay, *Chaos: Classical and Quantum*. Niels Bohr Inst., Copenhagen, 2016.
- [48] S. A. Hartnoll, *Wheeler-DeWitt states of the AdS-Schwarzschild interior*, *JHEP* **01** (2023) 066, [[2208.04348](#)].
- [49] G. W. Gibbons, S. W. Hawking and M. J. Perry, *Path Integrals and the Indefiniteness of the Gravitational Action*, *Nucl. Phys. B* **138** (1978) 141–150.
- [50] J. Louko, *A Feynman Prescription for the Hartle-hawking Proposal*, *Class. Quant. Grav.* **5** (1988) L181.
- [51] A. Dasgupta and R. Loll, *A Proper time cure for the conformal sickness in quantum gravity*, *Nucl. Phys. B* **606** (2001) 357–379, [[hep-th/0103186](#)].
- [52] D. Marolf, *Gravitational thermodynamics without the conformal factor problem: partition functions and Euclidean saddles from Lorentzian path integrals*, *JHEP* **07** (2022) 108, [[2203.07421](#)].
- [53] B. Banihashemi and T. Jacobson, *On the lapse contour in the gravitational path integral*, *Phys. Rev. D* **111** (2025) 066014, [[2405.10307](#)].
- [54] M. Kontsevich and G. Segal, *Wick Rotation and the Positivity of Energy in Quantum Field Theory*, *Quart. J. Math. Oxford Ser.* **72** (2021) 673–699, [[2105.10161](#)].
- [55] E. Witten, *A Note On Complex Spacetime Metrics*, [2111.06514](#).
- [56] V. Moncrief, *Spacetime symmetries and linearization stability of the Einstein equations. I*, *J. Math. Phys.* **16** (1975) 493.
- [57] D. Marolf and I. A. Morrison, *Group Averaging for de Sitter free fields*, *Class. Quant. Grav.* **26** (2009) 235003, [[0810.5163](#)].
- [58] J. Cotler and K. Jensen, *Isometric Evolution in de Sitter Quantum Gravity*, *Phys. Rev. Lett.* **131** (2023) 211601, [[2302.06603](#)].
- [59] V. Balasubramanian, A. Lawrence, J. M. Magan and M. Sasieta, *Microscopic Origin of the Entropy of Black Holes in General Relativity*, *Phys. Rev. X* **14** (2024) 011024, [[2212.02447](#)].
- [60] H. Maxfield, *Counting states in a model of replica wormholes*, [2311.05703](#).
- [61] J. M. Maldacena and L. Maoz, *Wormholes in AdS*, *JHEP* **02** (2004) 053, [[hep-th/0401024](#)].
- [62] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, *JHEP* **08** (2020) 044, [[2002.08950](#)].

- [63] J. Louko and R. D. Sorkin, *Complex actions in two-dimensional topology change*, *Class. Quant. Grav.* **14** (1997) 179–204, [[gr-qc/9511023](#)].
- [64] X. Dong, A. Lewkowycz and M. Rangamani, *Deriving covariant holographic entanglement*, *JHEP* **11** (2016) 028, [[1607.07506](#)].
- [65] D. Marolf and H. Maxfield, *The page curve and baby universes*, *Int. J. Mod. Phys. D* **30** (2021) 2142027, [[2105.12211](#)].
- [66] S. Colin-Ellerin, X. Dong, D. Marolf, M. Rangamani and Z. Wang, *Real-time gravitational replicas: Formalism and a variational principle*, *JHEP* **05** (2021) 117, [[2012.00828](#)].
- [67] B. Dittrich, T. Jacobson and J. Padua-Argüelles, *de Sitter horizon entropy from a simplicial Lorentzian path integral*, *Phys. Rev. D* **110** (2024) 046006, [[2403.02119](#)].