Local subsystems in gauge theory and gravity

William Donnelly[a] **and Laurent Freidel**[b]

[a]*Department of Physics, University of California, Santa Barbara,*
Santa Barbara, California 93106, U.S.A.

[b]*Perimeter Institute for Theoretical Physics,*
31 Caroline St. N, N2L 2Y5, Waterloo ON, Canada

E-mail: donnelly@physics.ucsb.edu, lfreidel@perimeterinstitute.ca

Abstract: We consider the problem of defining localized subsystems in gauge theory and gravity. Such systems are associated to spacelike hypersurfaces with boundaries and provide the natural setting for studying entanglement entropy of localized subsystems. We present a general formalism to associate a gauge-invariant classical phase space to a spatial slice with boundary by introducing new degrees of freedom on the boundary. In Yang-Mills theory the new degrees of freedom are a choice of gauge on the boundary, transformations of which are generated by the normal component of the nonabelian electric field. In general relativity the new degrees of freedom are the location of a codimension-2 surface and a choice of conformal normal frame. These degrees of freedom transform under a group of *surface symmetries*, consisting of diffeomorphisms of the codimension-2 boundary, and position-dependent linear deformations of its normal plane. We find the observables which generate these symmetries, consisting of the conformal normal metric and curvature of the normal connection. We discuss the implications for the problem of defining entanglement entropy in quantum gravity. Our work suggests that the Bekenstein-Hawking entropy may arise from the different ways of gluing together two partial Cauchy surfaces at a cross-section of the horizon.

Contents

1   Introduction

Entanglement plays an increasingly prominent role in fundamental physics. It has long been recognized that the entanglement naturally present in the vacuum could account for the Bekenstein-Hawking entropy of black holes [1–3] (see [4] for a review). In fact it is conjectured that the scaling of entanglement entropy with area is a universal feature of near-vacuum states in quantum gravity [5]. More ambitiously, it has been proposed that the geometry of spacetime [6] and its topology [7] may emerge from entanglement. Moreover, if the Bekenstein-Hawking entropy really is entanglement entropy, then all regions of space should possess an entropy and therefore be subject to the laws of thermodynamics. Applying the laws of thermodynamics to local Rindler horizons [8], or to small spherical regions of spacetime [9], one can derive the Einstein equation as an equation of

state. Together these results point toward idea that entanglement, an essentially quantum phenomenon, is deeply connected to classical spacetime geometry and its dynamics.

Despite the central role that entanglement is conjectured to play in quantum gravity, there is so far no sharp definition of entanglement between regions of space that would apply in quantum gravity. In order to define entanglement entropy, we require that for any Cauchy surface divided into two disjoint regions Σ and $\overline{\Sigma}$, there is a tensor product factorization of the total Hilbert space

$$\mathcal{H} = \mathcal{H}_\Sigma \otimes \mathcal{H}_{\overline{\Sigma}}. \tag{1.1}$$

We can then define for any density matrix ρ, a reduced density matrix $\rho_\Sigma := \mathrm{Tr}_{\overline{\Sigma}}\rho$ and an entanglement entropy

$$S(\rho_\Sigma) = -\mathrm{Tr}\rho_\Sigma \log \rho_\Sigma. \tag{1.2}$$

We do not expect the Hilbert space of quantum gravity to be naturally equipped with a tensor product structure such as (1.1). An essential property of the tensor product is that operators within the different factors commute: $[A \otimes I, I \otimes B] = 0$ for any operators A and B. There are strong arguments that such locally supported commuting operators simply cannot exist in quantum gravity [10, 11]. The basic obstruction is diffeomorphism invariance: physical operators in quantum gravity should be diffeomorphism invariant, but the action of a diffeomorphism is to permute the points of spacetime. Thus an invariant operator with support at any spacetime point must also have support on every other spacetime point. Whatever notion of local subsystems survives in quantum gravity must clearly be a generalization of (1.1). The goal of this paper is to unravel the mechanism behind such a generalisation and the basis for its construction.

The absence of tensor factorisation is expected for any theory with gauge symmetry. In a gauge theory the initial data must obey constraints, which take the form of differential equations on each time slice. In Yang-Mills theory the constraint is Gauss' law, and in gravity it is the diffeomorphism constraints (often further divided into a spatial diffeomorphism constraint and a timelike Hamiltonian constraint). Because of these constraint equations, initial data in the two regions Σ and $\overline{\Sigma}$ cannot be specified independently. This fact is reflected at the quantum level by the failure of the factorisation (1.1).

Before proceeding to describe our construction, we discuss some alternative perspectives on the problem of entanglement entropy and localized subsystems in field theory. The most straightforward way to define entanglement entropy for a field theory is to regularize on a lattice. There each site carries a Hilbert space, which could be for example that of a harmonic oscillator, or a finite-dimensional system. Then the Hilbert space of each region is simply the tensor product of all the sites inside the region, and the factorisation structure (1.1) is manifest. One can therefore define entanglement entropy for any quantum field theory that arises as a continuum limit of such a lattice system. Although the entanglement entropy defined in this way is divergent and depends on the precise lattice realization of the theory, certain universal quantities can be extracted from it. One might hope to define entanglement entropy in quantum gravity via such a bottom-up approach. Of course, this requires gravity to emerge from a theory with a local tensor product struc-

ture, as in string-net models [12]. It has recently been suggested that the bulk gravity theory in AdS/CFT must have an ultraviolet completion that does factor in this way [13].

In such a lattice model, gauge invariance is not fundamental. Instead, it emerges only within a low-energy sector of the Hilbert space. This low-energy sector does not admit a factorization into local subsystems; only the full ultraviolet completion does. Entanglement entropy of the low-energy gauge-invariant states is then defined by embedding them into a larger Hilbert space which does factorize. We will see that a similar embedding can be defined directly in the gauge theory, without specifying a lattice regularization.

An alternative approach is to identify a region of spacetime by its algebra of local observables, as is standard in the algebraic approach to QFT [14]. One can define the entropy of a region to be the entropy of its observable algebra, as advocated in ref. [15]. This approach to subsystems is problematic for several reasons [11]. First, in quantum field theory the algebra of observables associated to a spacetime region is a type III von Neumann algebra [16]. Unlike a type I von Neumann algebra, a type III von Neumann algebra is not associated with a factorization of the Hilbert space. Since the type III property depends on the ultraviolet behavior of quantum field theory, it is not clear whether this property will persist in quantum gravity. A much more serious issue is the absence of compactly supported gauge-invariant observables in gravity [10]. Thus it would seem that the algebraic approach, which is based on algebras of compactly supported gauge-invariant observables, does not know how to assign an entropy to systems without local gauge-invariant operators.

Instead we will follow an approach that has been called the extended Hilbert space construction [17–21]. We can divide the constraints of the theory into three classes: those with support in Σ, those with support in $\overline{\Sigma}$, and those with nontrivial support on the boundary. To the region Σ we assign an extended Hilbert space $\mathcal{H}_\Sigma$, consisting of states that satisfy the gauge condition away from the entangling surface $S = \partial\Sigma$, i.e. which are annihilated by the constraints with support only inside Σ. These states of the extended Hilbert spaces $\mathcal{H}_\Sigma$ and $\mathcal{H}_{\overline{\Sigma}}$ generally transform nontrivially under transformations with nontrivial support on the boundary, so the Hilbert space carries a nontrivial representation of the boundary gauge group. We can then identify the physical Hilbert space as a subspace of the tensor product

$$\mathcal{H} \subset \mathcal{H}_\Sigma \otimes \mathcal{H}_{\overline{\Sigma}}. \tag{1.3}$$

The space $\mathcal{H}$ consists of those states that are invariant under boundary gauge transformations. Equivalently, it is the space of states annihilated by the constraints with nontrivial support on the boundary.

The extended Hilbert space construction can be defined precisely in Yang-Mills theory on a lattice [18] and in two dimensions [22]. Here the extended Hilbert space $\mathcal{H}_\Sigma$ consists of Wilson loops (or more generally, generalized spin network states [18]) which are allowed to have open ends on the boundary S. This is in a sense the minimal extension of the Hilbert space: in order for $\mathcal{H}$ to contain gauge-invariant operators such as a Wilson loop in a representation R, the space $\mathcal{H}_\Sigma$ has to contain a state that looks like a Wilson loop in the interior of Σ, but which terminates on S. The endpoint of the Wilson line acts as a surface charge transforming in the representation R. The generators of this gauge transformation

are the nonabelian electric fields along links that cross the boundary. The subspace $\mathcal{H}$ consists of the gauge-invariant states for which the surface charges associated on one side of the boundary match up with the surface charges on the other side to form singlets. Equivalently, the normal component of the electric field must match across the boundary.

The result of this construction is that the gauge symmetry of the boundary is promoted to a physical symmetry on $\mathcal{H}_\Sigma$. The entanglement entropy that results from the embedding (1.3) includes a contribution from the surface charge degrees of freedom. While it may seem like overcounting to count gauge-related states in the entanglement entropy, multiple lines of evidence point to this being the physically relevant definition of entanglement entropy. For example, in ref. [22] it is shown that such states are necessary for consistency of the thermal description of the de Sitter vacuum. Additional evidence for this is provided by the calculation in ref. [23], where the entanglement entropy of a quark-antiquark pair was found to contain a contribution from surface charges where the nonabelian electric string pierces the entangling surface. In abelian gauge theory, the surface charges are needed to provide agreement with Euclidean methods, and to obtain agreement between the conformal anomaly and the universal logarithmic term in the entanglement entropy [24, 25]. Thus the surface charge degrees of freedom are an essential component of the horizon entropy.

Our goal is twofold. First, we want to revisit and improve the previous construction of the extended Hilbert space in gauge theory by analysing its classical analog, the extended phase space. Second, we would like to define an analogous extended phase space for gravity.

We first show how to assign to each region of space Σ an extended phase space $\mathcal{P}_\Sigma$, which contains additional degrees of freedom at the boundary. This extended phase space is the classical analog of the extended Hilbert space. Our analysis shows that gauge invariance dictates what type of additional degrees of freedom are needed and their commutation relations. These additional degrees of freedom are designed to make the phase space invariant under all gauge transformations of the original system, whether or not they vanish at the boundary. In addition to the gauge symmetries extended to the boundary, we discover that the boundary degrees of freedom transform covariantly under a new group of physical symmetries, which we call *surface symmetries*, that are implemented as Hamiltonian transformations on the extended phase space. The appearance of a new type of boundary symmetry group G_S where $S = \partial\Sigma$ as a result of gauge invariance is one of the key results of this paper. We find, quite remarkably, that the boundary symmetry group commutes with the gauge transformations; this implies that the charges generating G_S are gauge-invariant boundary observables. The second goal of this paper is to identify these charges for Yang-Mills and gravity. At the quantum level this implies that the extended Hilbert space $\mathcal{H}_\Sigma$ carries a representation of G_S.

We can use this Hamiltonian group action to define the the phase space associated to a union of regions $\Sigma \cup \overline{\Sigma}$. It is given by a *fusion* product

$$\mathcal{P}_{\Sigma \cup \overline{\Sigma}} = \mathcal{P}_\Sigma \times_{G_S} \mathcal{P}_{\overline{\Sigma}}, \tag{1.4}$$

where we take the Cartesian product of the phase spaces (analogous to the tensor product of Hilbert spaces) and mod out by the action of the surface symmetry group. At the

quantum level this construction is implemented by an *entangling* product

$$\mathcal{H}_{\Sigma \cup \overline{\Sigma}} = \mathcal{H}_\Sigma \otimes_{G_S} \mathcal{H}_{\overline{\Sigma}}, \tag{1.5}$$

where the entangling product $\otimes_{G_S}$ is a tensor product of modules: H_Σ acts as a right module for G_S and $H_{\overline{\Sigma}}$ acts as left module for G_S and we make the identification $(g\,|\psi\rangle\,;g\,|\overline{\psi}\rangle) \simeq (|\psi\rangle\,;|\overline{\psi}\rangle)$, for $g \in G_S$. In other words the entangling product restricts to states that are singlets under the diagonal action of surface symmetries.

We begin by constructing an extended phase space for Yang-Mills theory with a general gauge group G in section 2. We follow the covariant canonical formalism, which has been developed by many authors, and presented e.g. in [26]. The key step is to include the choice of gauge as a dynamical variable in the phase space.[1] In the mathematical language this is a choice of bundle trivialization, but locally it is simply a G-valued function $\varphi : S \to G$. Following the standard algorithm we define a presymplectic potential form θ, which can be integrated over a Cauchy surface Σ to define the presymplectic potential of the theory. If the Cauchy surface Σ has a boundary, the presymplectic potential obtained from this procedure is not gauge-invariant. We show how it can be made invariant by adding a boundary term that depends on φ and its conjugate momentum: the non-abelian surface charge E, which is a Lie algebra valued density. When this is done we obtain a presymplectic potential that is fully invariant under gauge transformations, including those which act at the boundary.

The introduction of the additional variable φ allows us to introduce a new symmetry of the system, which rotates φ without changing the other fields. Thus we distinguish between two different types of transformations:

$$\text{Gauge symmetry:} \qquad \delta_\alpha \varphi = \varphi \alpha, \qquad\qquad \delta_\alpha A = \mathrm{d}_A \alpha \tag{1.6}$$

$$\text{Surface symmetry:} \qquad \Delta_\alpha \varphi = -\alpha \varphi, \qquad\qquad \Delta_\alpha A = 0 \tag{1.7}$$

The former is pure gauge, acts on the right of the trivialisation, and has vanishing generator $H[\alpha]$. This generator can be expressed as a bulk term proportional to the Gauss constraint, and a boundary term that enforces the identification of the normal electric field with the surface charge (see section 2.5). The latter is a symmetry, acts on the left of the trivialisation, and its generator $E[\alpha]$ is nonvanishing and is given by the the non-abelian surface charge. The vanishing of the boundary component of the gauge generator $H[\alpha]$ can be understood as a boundary condition that identifies the non-abelian surface charge density E with the normal component of the electric field in the frame φ. Explicitly this reads

$$n^a s^b F_{ab} = (\varphi^{-1} E \varphi), \tag{1.8}$$

n^a is the unit time normal to the slice Σ and s^b is the unit normal of S within Σ, F is the Yang-Mills curvature. These normal electric fields therefore generate an algebra of

---

[1]This procedure is similar to that followed in ref. [27], who introduced an asymptotic coordinate system as additional dynamical variables to the phase space of asymptotically flat gravity in order to have a phase space carrying a representation of the Poincarë group.

non-abelian surface charges. The Poisson brackets of this extended phase space coincide with the commutators of the extended Hilbert space of lattice gauge theory; we have thus succeeded in constructing a continuum version of the lattice construction [18].

The most significant advantage of our covariant formalism is that it extends naturally to general relativity, as we show in section 3. Here in addition to a metric g_{ab}, we introduce a variable $X^a : \mathbb{R}^D \to M$ that specifies the coordinate location of the entangling surface $S = X(s)$, where s is a fixed $(D-2)$-dimensional closed surface in $\mathbb{R}^D$. X^a plays a role analogous to φ which specifies the choice of boundary gauge in Yang-Mills. The non-temporal variable X^a drops out of the presymplectic potential when integrated over a manifold without boundary. As shown in [28] the temporal component δT does not in general, unless we restrict to hypersurface normal deformations. On a surface Σ with boundary, we show that the addition of an X^a-dependent boundary term makes the presymplectic potential diffeomorphism-invariant, even for diffeomorphisms supported on the boundary.

The additional phase space variable X^a again leads to physical symmetries of the extended phase space. Thus again we distinguish between two types of transformations:

$$\text{Gauge symmetry:} \qquad \delta_V X^a = -V^a(X), \qquad \delta_V g_{ab} = \mathcal{L}_V g_{ab} \qquad (1.9)$$

$$\text{Surface symmetry:} \qquad \Delta_w X^a = (\partial_b X^a) w^b, \qquad \Delta_w g_{ab} = 0 \qquad (1.10)$$

V is a vector field on M that labels infinitesimal diffeomorphism of spacetime: these are gauge. w is a vector field on $\mathbb{R}^D$ that generates an infinitesimal change of the surface S: these are the surface symmetries. The gravitational case is more subtle than Yang-Mills because of the different ways that a transformation of X^a can act in a neighbourhood of the codimension-2 surface S. Thus we distinguish between three classes of surface transformations:

- *Surface boosts* linearly transform the normal plane of S in a position-dependent way.

- *Surface diffeomorphisms* are reparametrizations of S that move S tangent to itself.

- *Surface translations* move the surface S in the normal direction.

Previous works [29, 30] and especially [31–33] (see [34] for a more recent overview) have focused on the surface translations. In order to implement the surface translations as canonical transformations, boundary conditions are usually imposed at the surface S, and the correct way to choose those boundary conditions remains somewhat mysterious. Here we focus instead on the first two classes of transformations, which we call the *surface-preserving symmetries*. These are the boundary symmetries that preserve the location of the surface as a whole. We find canonical generators for these symmetries, without the need to specify the location of the surface or any associated boundary conditions. However it is clear that the full story must include both classes of transformations, especially if we are to consider dynamically evolving entangling surfaces. While we hope that our formalism will lend insight to the quantization of the surface translations, we leave this task for future work.

The central result of this paper is twofold. First, we identify the surface-preserving symmetry group which is given by the semi-direct product

$$G_S = \text{Diff}(S) \ltimes \text{SL}(2, \mathbb{R})^S \tag{1.11}$$

of diffeomorphisms of the entangling surface, times the algebra of surface boosts represented as group valued maps $S \to \text{SL}(2, \mathbb{R})$. We also show that the action of this symmetry group is Hamiltonian on the extended phase space of gravity and identify the Hamiltonian generators in terms of intrinsic and extrinsic geometry of S. These Hamiltonian generators play the role of gravitational surface charges for the entangling surface S. They are the analog of the dressed normal electric field (1.8) for Yang-Mills. Remarkably they can be described explicitly as follows.

To evaluate the Hamiltonian in terms of components of the metric we introduce a $2 + (D - 2)$ decomposition and adapted coordinates (x^i, σ^μ). σ^μ are coordinates on S, and $(x^i)_{i=0,1}$ are coordinates in the two normal directions. In these coordinates we can parametrize the metric as

$$ds^2 = h_{ij}dx^i dx^j + q_{\mu\nu} \left(d\sigma^\mu - A_i^\mu dx^i \right) \left(d\sigma^\nu - A_j^\nu dx^j \right). \tag{1.12}$$

Here $q_{\mu\nu}$ is the induced metric on S, h_{ij} a generalized lapse which defines the normal geometry and A_i^μ a generalized shift. A_i^μ can be viewed as a connection on the normal bundle of S. The generators of diffeomorphisms and $\text{SL}(2, \mathbb{R})$ symmetries are integrals of surface charge densities, given by the normal curvature and the conformal normal geometry, respectively:

$$\sqrt{q} \left(\partial_0 A_1^\mu - \partial_1 A_0^\mu + [A_0, A_1]^\mu \right)/\sqrt{h}, \qquad -\sqrt{q} \, h_{jk} \epsilon^{ki}/\sqrt{h}. \tag{1.13}$$

Here $\sqrt{q}$ is the volume form on S, while $\sqrt{h}$ is the square root of the determinant of the normal metric, and ϵ^{ij} is the two-dimensional Levi-Civita symbol.

We see that the algebra of surface boosts is generated by conformal components of the metric normal to S and form a local $\text{SL}(2, \mathbb{R})$ algebra whose Casimir is the area element of the surface S. The diffeomorphisms are generated by the curvature of the normal connection, which vanishes if the normal planes are integrable. This is our main result. This result, along with the identification of the fusion product of phase space and the entangling product, form the basis of understanding gauge systems with boundaries.

While quantization is beyond the scope of this work, our approach shows that the Hilbert space associated with Σ must carry a unitary representation of the boundary symmetry group G_S. In section 4 we mention some general results in representation theory which may be useful. A key point is the identification of the area element of the surface with the Casimir of $\text{SL}(2, \mathbb{R})$. The representation theory of this algebra therefore singles out the area of the entangling surface S as the physical quantity that labels the dimension of the irreducible representations, and area preserving diffeomorphisms as a key subgroup to represent unitarily. We conclude in section 5 with many areas for future work. Some technical details of our formalism have been exiled to two appendices.

2 Yang-Mills symplectic potential

We first construct an extended phase space for bounded regions in Yang-Mills theory. In addition to being of interest in its own right, Yang-Mills theory provides a useful warm-up exercise for gravity and allows us to review some aspects of the covariant canonical formalism. Moreover, it will allow us to see that the method we have developed is a continuum generalization of the extended Hilbert space approach to lattice Yang-Mills theory [17–21].

Consider Yang-Mills theory with gauge group G. The field variable is a connection A, which can be represented in a trivialization as a $\mathfrak{g}$-valued 1-form. The Lagrangian density is given by the D-form

$$L = -\frac{1}{2}\text{Tr}[F \wedge \star F]. \tag{2.1}$$

where $\star$ denotes the D-dimensional Hodge duality operation[2] and $F = \mathrm{d}A + \frac{1}{2}[A, A]$ is the curvature two-form.

The first step of deriving the canonical formalism is to find the symplectic potential of the theory, which we do following the method of refs. [26, 35–38]. The space of classical solutions of a Lagrangian system is naturally equipped with a presymplectic structure. The symplectic potential itself is obtained from an on-shell variation of the action. We introduce a differential δ on field space which satisfies Leibniz rule and squares to zero: $\delta^2 = 0$. It acts on the space of solutions of the Lagrangian system mapping infinitesimally solutions to solutions and it is such that the product of field differentials anti commute $\delta A \delta B = -\delta B \delta A$ for A, B two scalar fields. Following the conventions of ref. [36] we do not include an explicit wedge product between the 1-forms δ_g; we reserve the notation $\wedge$ for spacetime forms and keep the antisymmetrization of δ implicit. This field space differential δ should not be confused with the spacetime differential denoted d. A general field is a (p, q) form where p denotes the degree of the form in field space and q the degree as a spacetime form.

We begin by computing the symplectic potential current density θ by varying the Lagrangian,

$$\delta L[A] = \text{Tr}[\delta A \wedge (\mathrm{d}_A \star F)] + \mathrm{d}\theta[A, \delta A]. \tag{2.2}$$

The first term gives the equation of motion, $\mathrm{d}_A(\star F) := \mathrm{d}(\star F) + [A, \star F] = 0$. The second term yields the symplectic potential current density:

$$\theta[A, \delta A] = -\text{Tr}[\delta A \wedge \star F]. \tag{2.3}$$

θ depends on the field A and is linear in the variation δA; it is a $(1, D - 1)$-form, a one-form in field space in the sense of ref. [36] and a $(D - 1)$-form in spacetime. To obtain the presymplectic potential of the theory, we integrate θ over a Cauchy surface. The expression δA is analogous to the finite-dimensional expression dq where q is a canonical coordinate, and the integral of the current θ is analogous to the finite-dimensional symplectic potential $\sum_i p_i \mathrm{d}q^i$.

---

[2]Our convention is that $(\alpha \wedge \star \beta) = g(\alpha, \beta)\epsilon$ where ϵ is the volume form.

It is important to note that θ depends explicitly on A, and is not gauge-invariant. This is potentially a problem, because we would like to define θ as a one-form on the gauge orbits, and this is only possible if θ is gauge-invariant. As we will see this poses no problem when θ is integrated over a closed surface Σ without boundary, but when Σ has a boundary the symplectic structure does change nontrivially under gauge transformations. We will then show how to correct for this non-invariance by addition of a boundary term to the symplectic structure.

Let g^* denote the gauge transformation by g, which acts on A and F as

$$A \to g^*(A) = gAg^{-1} - (\mathrm{d}g)g^{-1}, \tag{2.4}$$

$$F \to g^*(F) = gFg^{-1}. \tag{2.5}$$

Now consider how θ changes under this gauge transformation. After transformation the new symplectic potential current density is given by

$$
\begin{aligned}
\theta[g^*(A), \delta g^*(A)] &= -\mathrm{Tr}[(\delta(gAg^{-1}) - \delta(\mathrm{d}gg^{-1})) \wedge \star gFg^{-1}] \\
&= -\mathrm{Tr}[\big(\delta A - \mathrm{d}(g^{-1}\delta g) - [A, g^{-1}\delta g]\big) \wedge \star F].
\end{aligned}
\tag{2.6}
$$

We now introduce some notation that will clarify this calculation and generalize more readily to the gravitational case. The quantity δg is a somewhat unnatural object: it is an element of the tangent space to G at the point g. A more natural object is obtained by translating it back to a neighbourhood of the identity:

$$\delta_g := g^{-1}\delta g. \tag{2.7}$$

δ_g is an element of the Lie algebra, and the transformation of the presymplectic potential form (2.6) can be expressed in terms of the gauge covariant derivative of δ_g:

$$\mathrm{d}_A(\delta_g) = \mathrm{d}(g^{-1}\delta g) + [A, g^{-1}\delta g]. \tag{2.8}$$

Thus we have

$$\theta[g^*A, \delta g^*A] = -\mathrm{Tr}[(\delta A - \mathrm{d}_A(\delta_g)) \wedge \star F]. \tag{2.9}$$

The presence of the second term in eq. (2.9) shows precisely the degree to which the symplectic potential current density is not gauge invariant.

The reason that θ is not gauge invariant is that the variation δ does not commute with the gauge transformation g^*. Instead the difference can be expressed as an infinitesimal gauge transformation by δ_g:

$$\delta g^*(A) = g^*(\delta A - \mathrm{d}_A \delta_g). \tag{2.10}$$

Note that δA being a difference of two connections transforms as $g^*(\delta A) = g\, \delta A\, g^{-1}$. Similarly, for the field strength F we have

$$\delta g^*(F) = g^*(\delta F + [\delta_g, F]), \tag{2.11}$$

which is again the transformation rule for F under an infinitesimal gauge transformation by δ_g.

Later we will need the variation of δ_g itself, which is given by

$$\delta(\delta_g) = -\tfrac{1}{2}[\delta_g, \delta_g]. \tag{2.12}$$

Recall that the two δ's in this formula correspond to two different (anticommuting) variations.

To obtain the presymplectic potential Θ of the theory, we integrate the symplectic potential current over a Cauchy surface Σ:

$$\Theta_\Sigma[A, \delta A] = \int_\Sigma \theta[A, \delta A] = - \int_\Sigma \mathrm{Tr}\,[\delta A \wedge \star F]. \tag{2.13}$$

This is not manifestly gauge invariant because of the nontrivial transformation law for θ in eq. (2.9). But the latter term can be integrated by parts, using the equation of motion $\mathrm{d}_A \star F \doteq 0$, with $\doteq$ denoting equality on shell. We find that the presymplectic potential changes by a boundary term,

$$\Theta_\Sigma[g^*A, \delta g^*A] \doteq \Theta_\Sigma[A, \delta A] + \oint_S \mathrm{Tr}[\star F \delta_g]. \tag{2.14}$$

When θ is integrated over a Cauchy surface without boundary, or when the boundary conditions are such that the boundary term in (2.14) vanishes, the presymplectic potential is gauge invariant. However we are interested in associated a phase space with a region Σ whose boundary S is the entangling surface. In this case we do not wish to impose any particular boundary conditions at S, since the entangling surface is not a physically distinguished location as far as the gauge field is concerned.

In order to define a gauge-invariant symplectic potential we need to add to Θ_Σ a term whose gauge transformation compensates that in eq. (2.14). To construct such a potential we recall that a connection on a bundle is not specified just by the 1-form A but also by a local trivialization φ, which we may think of, at least locally, as a G-valued function $\varphi : S \to G$. Under a gauge transformation φ transforms as $g^*(\varphi) = \varphi g^{-1}$. We can now augment the symplectic structure with the term

$$\Theta_S[A, \varphi, \delta\varphi] = \oint_S \mathrm{Tr}[\star F \delta_\varphi] \tag{2.15}$$

where analogous to δ_g we define $\delta_\varphi := \varphi^{-1}\delta\varphi$. Under a gauge transformation δ_φ transforms as

$$\delta_{g^*(\varphi)} = g^*(\delta_\varphi - \delta_g) \tag{2.16}$$

which leads to the gauge transformation law for Θ_S

$$\Theta_S[g^*A, g^*\varphi, \delta g^*\varphi] = \Theta_S[F, \varphi, \delta\varphi] - \oint_S \mathrm{Tr}[\star F \delta_g]. \tag{2.17}$$

The second term in (2.17) is exactly what is needed to cancel the second term in (2.14). Combining these two terms we obtain the presymplectic potential

$$\Theta[A, \delta A, \varphi, \delta\varphi] = \Theta_\Sigma[A, \delta A] + \Theta_S[A, \varphi, \delta\varphi] \tag{2.18}$$

$$= - \int_\Sigma \mathrm{Tr}[\delta A \wedge \star F] + \oint_S \mathrm{Tr}[\star F \delta_\varphi]. \tag{2.19}$$

This presymplectic potential is gauge invariant for any spacelike surface Σ, even for gauge transformations with support on the boundary.

2.1 Gauge invariance

The above presymplectic potential satisfies, by construction, the gauge invariance property

$$\Theta[g^*A, \delta g^*A, g^*\varphi, \delta g^*\varphi] = \Theta[A, \delta A, \varphi, \delta\varphi]. \tag{2.20}$$

Note that this depends not only on the gauge field A everywhere on Σ (including the boundary S), but also on the trivialization φ on S. Because Θ is gauge-invariant, we can pull it back to define a symplectic potential on the space gauge orbits, which is the true physical phase space of the theory. This means that for any α valued in the Lie algebra $\mathfrak{g}$, the infinitesimal gauge transformation

$$\delta_\alpha \varphi = \varphi\alpha, \qquad \delta_\alpha A = \mathrm{d}_A\alpha, \qquad \delta_\alpha F = [F, \alpha]. \tag{2.21}$$

is a null direction of the total symplectic form, for any α including those with support on the boundary. We will now show explicitly that the canonical generator for this transformation vanishes on-shell. To find the generators of these transformations, we take the variation of the symplectic potential to obtain the presymplectic variational 2-form $\Omega = \delta\Theta$. This presymplectic form is the sum of a bulk term Ω_Σ and a boundary term Ω_S, where the latter depends on the boundary degree of freedom φ:

$$\Omega_\Sigma = \int_\Sigma \mathrm{Tr}[\delta A \wedge \star\delta F], \tag{2.22}$$

$$\Omega_S = \oint_S \mathrm{Tr}[\star\delta F \delta_\varphi - \star F \delta_\varphi \delta_\varphi]. \tag{2.23}$$

We are now interested in constructing the generators of gauge transformations associated with the bulk and boundary symplectic structure in order to determine the Poisson brackets between boundary observables. Recall that a Hamiltonian H determines a variational vector field δ_H via

$$\Big\{H, f\Big\} = \delta_H f. \tag{2.24}$$

Since we have already introduced a field differential δ it is natural to also introduce a field "interior product": given any infinitesimal field transformation $\delta_\alpha A$, we denote I_{δ_α} the inner product along this variation that maps n-form in field space to $(n-1)$-forms. I_{δ_α} acts trivially on functionals of A (field space 0-forms), satisfies the graded Leibniz rule (see appendix A) and is defined on field one-forms by

$$I_{\delta_\alpha}\delta A := \delta_\alpha A. \tag{2.25}$$

The Poisson bracket (2.24) can then be expressed in terms of the interior product and the symplectic structure Ω via

$$I_{\delta_F} I_{\delta_H} \Omega = [I_{\delta_H}\Omega](\delta_F) = \Omega(\delta_H, \delta_F) = \Big\{F, H\Big\}. \tag{2.26}$$

where I is the interior product. Thus the Hamiltonian H and the variational vector field it generates are related by

$$I_{\delta_H}\Omega = \Omega(\delta_H, \cdot) = \delta H. \tag{2.27}$$

We can use the relation (2.27) to find the generator of gauge transformations. Let α be a Lie-algebra-valued function on Σ and δ_α denote the gauge transformation (2.21). We introduce the bulk and boundary Hamiltonian generators $H_\Sigma[\alpha]$ and $H_S[\alpha]$ respectively such that $I_{\delta_\alpha}\Omega_\Sigma = \delta H_\Sigma[\alpha]$, and $I_{\delta_\alpha}\Omega_S = \delta H_S[\alpha]$. One finds that they are given by

$$H_\Sigma[\alpha] = \int_\Sigma \text{Tr}[\text{d}_A\alpha \wedge \star F], \qquad H_S[\alpha] = -\int_S \text{Tr}[\star F\alpha]. \tag{2.28}$$

For instance we can explicitly check the boundary variation

$$\delta H_S[\alpha] = -\oint_S \star\text{Tr}[\delta F\,\alpha] = -\oint_S \star\text{Tr}[\delta F(\varphi^{-1}\delta_\alpha\varphi) + (-\delta_\alpha F + [F, \varphi^{-1}\delta_\alpha\varphi])\varphi^{-1}\delta\varphi] = I_{\delta_\alpha}\Omega_S.$$

We have used that the second term vanishes since $(\varphi^{-1}\delta_\alpha\varphi) = \alpha$. We can now check that the full generator $H[\alpha] = H_\Sigma[\alpha] + H_S[\alpha]$ vanishes on-shell:

$$H_\Sigma[\alpha] + H_S[\alpha] = -\int_\Sigma \text{Tr}[\alpha(\text{d}_A \star F)] \;\hat{=}\; 0. \tag{2.29}$$

This is the expression of the gauge invariance of our formulation.

2.2 Formal derivation of gauge invariance

Since we would like to generalize our results to the case of gravity, and since direct computations in gravity such as the one done in the previous section are very cumbersome, it is convenient to develop a formal but simpler proof of the gauge invariance of the presymplectic potential. Such a proof will then generalize easily to the gravity case and also be conceptually clearer. Moreover, this proof elucidates the key mechanisms behind the restoration of gauge invariance by the addition of boundary degrees of freedom.

Using the field interior product we can define a field Lie derivative[3]

$$L_{\delta_\alpha} := \delta I_{\delta_\alpha} + I_{\delta_\alpha}\delta. \tag{2.30}$$

We can now evaluate

$$I_{\delta_\alpha}\Omega = L_{\delta_\alpha}\Theta - \delta(I_{\delta_\alpha}\Theta). \tag{2.31}$$

The symplectic potential we have constructed is invariant under gauge transformations, which was the main motivation to introduce the trivialisation degree of freedom φ. Indeed the infinitesimal version of (2.20) when $g = 1 + \alpha + O(\alpha^2)$ reads

$$L_{\delta_\alpha}\Theta = 0. \tag{2.32}$$

This shows that the Hamiltonian is simply given by $H[\alpha] = -I_{\delta_\alpha}\Theta$. This can be evaluated directly

$$I_{\delta_\alpha}\Theta = -\int_\Sigma \text{Tr}[\delta_\alpha A \wedge \star F] + \oint_S \text{Tr}[\star F\alpha] = \int_\Sigma \text{Tr}[\alpha(\text{d}_A \star F)] \;\hat{=}\; 0. \tag{2.33}$$

Here we recognize the contraction of the Yang-Mills equation of motion with the gauge parameter α pulled back to Σ. As we will see the same derivation applies in the gravity case.

---

[3]L and its properties are described in more detail in appendix A.

2.3 Boundary observables

The addition of φ to the phase space allows us to construct new gauge-invariant operators. These are operators that depend on the choice of trivialization on the boundary. One is the electric field normal to the boundary, $\varphi \star F \varphi^{-1}$. Another is a Wilson line that begins and ends on the boundary,

$$W(\gamma) = \varphi(x_1) \, \mathcal{P} \exp\left(i \int_\gamma A \right) \, \varphi(x_2)^{-1} \tag{2.34}$$

where γ is a curve with endpoints x_1 and x_2 on the boundary.

Let us now focus on the gauge invariant normal electric field observable:

$$E[\alpha] := \oint_S \mathrm{Tr}[\star \varphi F \varphi^{-1} \alpha]. \tag{2.35}$$

Its variation is given by

$$\delta E[\alpha] = \oint_S \mathrm{Tr}\left[\star \left(\delta F + [\delta_\varphi, F] \right) \varphi^{-1} \alpha \varphi \right] = I_{\Delta_\alpha} \Omega \tag{2.36}$$

where we define the variational vector field Δ_α such that $I_{\Delta_\alpha} \delta_\varphi = -\varphi^{-1} \alpha \varphi$. Δ_α acts only on the boundary trivialisation, so we have

$$\Delta_\alpha \varphi = -\alpha \varphi, \qquad \Delta_\alpha A = 0. \tag{2.37}$$

$E[\alpha]$ therefore generates a local rotation of the boundary trivialisation. This rotation acts on the left of the trivialisation, unlike the gauge transformation which acts on the right. This equation is equivalent to the statement that the Poisson bracket of $\{E[\alpha], \cdot\} = \Delta_\alpha$ with a boundary observable generates the transformation Δ_α. Accordingly the previous equation determines the Poisson brackets between E_α's via

$$\{E[\alpha], E[\beta]\} = \Delta_\alpha E[\beta] = - \oint_S \star \mathrm{Tr}[\beta[\alpha, \varphi F \varphi^{-1}]] = E[[\alpha, \beta]]. \tag{2.38}$$

Thus we find that the different components of $E := \star(\varphi F \varphi^{-1})$ pulled back on S do not commute, but satisfy the current algebra of the group G. Expressed in terms of a set of generators, $E = E^a \tau_a$ with $[\tau_a, \tau_b] = f_{ab}{}^c \tau_c$ and $\mathrm{Tr}[\tau_a \tau_b] = \delta_{ab}$, we can expand (2.38) to obtain the explicit commutation relations

$$\{E^a(x), E^b(y)\} = \delta(x - y) f^{ab}{}_c E^c(x), \tag{2.39}$$

$$\{E^a(x), \varphi(y)\} = -\delta(x - y) \tau^a \varphi(x), \tag{2.40}$$

$$\{\varphi(x), \varphi(y)\} = 0. \tag{2.41}$$

This is the current algebra of the group G, i.e. the algebra satisfied by nonabelian gauge charges on the surface S. Thus the dressed normal electric field E acts as a surface charge.

– 13 –

2.4 Gauge versus symmetry

As a summary, we see that the introduction of the boundary degree of freedom φ has two effects. First, it restores the gauge invariance of the theory. Second, it introduces a new boundary symmetry which rotates the boundary observables (φ, E). The gauge transformations acts on the bulk variables and on the *right* of the boundary trivialisation

$$\delta_\alpha \varphi = \varphi \alpha, \qquad \delta_\alpha A = \mathrm{d}_A \alpha. \tag{2.42}$$

Its Hamiltonian $H[\alpha] = \int_\Sigma \mathrm{Tr}[(\mathrm{d}_A \star F)\alpha]$ vanishes on-shell. The boundary symmetry leaves the bulk variables invariant and acts on the *left* of the boundary trivialisation

$$\Delta_\alpha \varphi = -\alpha \varphi, \qquad \Delta_\alpha A = 0. \tag{2.43}$$

with Hamiltonian $E[\alpha] = \oint_S \mathrm{Tr}[\star(\varphi F \varphi^{-1})\alpha]$. As a result, gauge transformations commute with symmetries. This implies that the symmetry generator $E[\alpha]$ is gauge invariant. At the quantum level the Hilbert space $\mathcal{H}_\Sigma$ is annihilated by the constraints $H[\alpha]$ and carries a nontrivial representation of the symmetry algebra generated by $E[\alpha]$.

2.5 Extended phase space and gluing

We are now able to define precisely the extended phase space $\mathcal{P}_\Sigma$ associated with an open region Σ and the gluing property of phase space associated with two regions Σ_L, Σ_R. A key notion in order to describe these two procedures is the concept of symplectic reduction. Consider a phase space $(\mathcal{P}, \Omega)$ that admits a group action: $G \times \mathcal{P} \to \mathcal{P}$. This action is Hamiltonian if it preserves the symplectic structure. For an infinitesimal transformation δ_V this translates to the condition

$$L_{\delta_V} \Omega = 0. \tag{2.44}$$

From the Cartan definition of L_{δ_V} and the closure of Ω we conclude that $\delta I_{\delta_V} \Omega = 0$, hence locally in field space we have $I_{\delta_V} \Omega = \delta H_V$. The action is said to be globally Hamiltonian when the identity $I_{\delta_V} \Omega = \delta H_V$ is valid globally on field space.[4] H is the Hamiltonian of the group action.

Given such a Hamiltonian group action we can define the symplectic quotient which is denoted by a double quotient labelled either by the group or by the set of Hamiltonians:

$$\mathcal{P}//G = \mathcal{P}//\{H\} \tag{2.45}$$

The procedure is well-known from the theory of constrained Hamiltonian systems [39]. First one considers the constraint subspace $C_H \subset \mathcal{P}$ on which the constraint $H = 0$ is imposed. C_H is naturally equipped with a presymplectic structure $\Omega_H = i_H^* \Omega$ obtained by the pullback of the embedding map $i_H : C_H \hookrightarrow \mathcal{P}$. To obtain a symplectic manifold we need to quotient out the action of the group G, so that the phase space is then given by C_H/G. In the more general case where we are given a family of Hamiltonians $\{H\}$ we

---

[4]The appearance of central charges c are due to global obstructions where $I_{\delta_V} \Omega = \delta H_V + c$.

only need to divide out C_H by the kernel $\mathrm{Ker}(\Omega_H)$, so that the phase space is $\mathcal{P}//\{H\} := C_H/\mathrm{Ker}(\Omega_H)$. The advantage of this description is that it treats first and second class constraints simultaneously. The vector fields in $\mathrm{Ker}(\Omega_H)$ are the infinitesimal generators of a group of transformations we denote G_H. The double quotient notation refers to the fact that the phase space is obtained by a double operation: first a *restriction* onto the kernel C_H of constraints, and then then a *quotient* of this set by the group G_H generated by the Hamiltonian vectors preserving C_H. The group G_H and the Hamiltonians H therefore play a dual role in this double quotient.

We can now use the symplectic reduction to define a classical analog of the entangling product. Suppose that we have two symplectic spaces $(\mathcal{P}_1, \Omega_1)$ and $(\mathcal{P}_2, \Omega_2)$ and consider the constraints $H_{12} := H_1 - H_2 = 0$, which generate the group of transformations G_{12}. We can then define the fusion product

$$\mathcal{P}_1 \times_{G_{12}} \mathcal{P}_2 := (\mathcal{P}_1 \times \mathcal{P}_2) // G_{12}. \tag{2.46}$$

This fusion product or classical entangling product identifies the two Hamiltonians $H_1 = H_2$ and divides out by their flow. Thus in the fusion product the phase space variables are those acting on the space in which $H_1 = H_2$ that are also invariant under its action: observables O such that $\{O, H_1 - H_2\} = 0$.

We can now describe precisely the extended phase space for Yang-Mills. This description formalises the construction that was first given in [40], in the context of first order gravity. One starts with a bulk phase space $(\hat{\mathcal{P}}_\Sigma, \hat{\Omega}_\Sigma)$, whose phase space variables are the pull-back of $(\star F, A)$ on Σ and a boundary phase space $(\hat{\mathcal{P}}_S, \hat{\Omega}_S)$ whose phase space variables are given by the pairs (E, φ) on S. Here E is a $(D-2)$-form on S valued in the Lie algebra and φ is a G-valued function on S. The symplectic structures are given by

$$\hat{\Omega}_\Sigma = \int_\Sigma \mathrm{Tr}(\delta A \wedge \star \delta F), \qquad \hat{\Omega}_S = \oint_S \mathrm{Tr}(\delta E \delta \varphi \varphi^{-1} + E \delta \varphi \varphi^{-1} \delta \varphi \varphi^{-1}). \tag{2.47}$$

On the entangling surface S we recognise the natural symplectic structure of the group cotangent bundle $T^*(G^S) \simeq (G \ltimes \mathfrak{g})^S$. The fusion of these symplectic structures is obtained by imposing the boundary condition

$$H_S[\alpha] := \oint_S \left[(\star F) - \varphi^{-1} E \varphi \right] \alpha = 0 \tag{2.48}$$

for $\alpha \in \mathfrak{g}^S$. The extended phase space associated with Σ with boundary $\partial \Sigma = S$ is simply the fusion product

$$\mathcal{P}_\Sigma := \hat{\mathcal{P}}_\Sigma \times_{H_S} \hat{\mathcal{P}}_S. \tag{2.49}$$

What is remarkable is the fact that the boundary condition (2.48) which determine the fusion product is in fact *dictated* by gauge invariance. Indeed as we have seen the generator of gauge invariance $H[\alpha]$ contains two contributions, one from the bulk and one from the boundary. The boundary part of the Hamiltonian generator is given by (2.48)

$$H[\alpha] = H_\Sigma[\alpha] + H_S[\alpha], \qquad H_\Sigma[\alpha] = -\int_\Sigma \mathrm{Tr}[(\mathrm{d}_A \star F)\alpha]. \tag{2.50}$$

The gluing constraints (2.48) can therefore be understood as the boundary component of the Gauss law and results from demanding that $H[\alpha] = 0$ which imposes gauge invariance. Note that in this language the generator of boundary symmetries G_S is simply

$$E[\alpha] = \oint_S \mathrm{Tr}[E\alpha]. \tag{2.51}$$

We can now describe precisely the gluing procedure of two regions Σ_L and Σ_R sharing a common boundary $S = \Sigma_L \cap \Sigma_R$. It is given by the fusion product associated with the constraints

$$E_L[\alpha] - E_R[\alpha] = 0. \tag{2.52}$$

This constraint generates a diagonal action of the boundary symmetry group on $\mathcal{P}_{\Sigma_L}$ and $\mathcal{P}_{\Sigma_R}$. The extended phase space associated with the region $\Sigma_L \cup \Sigma_R$ is simply given by the fusion product

$$\mathcal{P}_{\Sigma_L \cup \Sigma_R} := \mathcal{P}_{\Sigma_L} \times_{G_S} \mathcal{P}_{\Sigma_R} = \left(\hat{\mathcal{P}}_{\Sigma_L} \times_{H_{S_L}} \hat{\mathcal{P}}_{S_L} \right) \times_{G_S} \left(\hat{\mathcal{P}}_{S_R} \times_{H_{S_R}} \hat{\mathcal{P}}_{\Sigma_R} \right) \tag{2.53}$$

We have described the procedure of extension and gluing at the classical level in terms of phase space. This procedure has a direct, and in some ways simpler, analog at the quantum level. At the quantum level the phase space $\mathcal{P}_\Sigma$ is replaced with a Hilbert space $\mathcal{H}_\Sigma$. The quantum analog of the symplectic quotient $\mathcal{P}//G$ is simply the quotient $\mathcal{H}/G$ of the Hilbert space by the group action. A simple quotient is enough since demanding invariance under the action of G is equivalent to imposing the Hamiltonian constraint $H |\psi\rangle = 0$, since the Hamiltonian H generates the action of G. That is at the quantum level we have the equality $\mathrm{Ker}(H) = \mathcal{H}/G$, which makes the symplectic reduction a conceptually simpler operation.[5] Accordingly, the extended Hilbert space is simply given by

$$\mathcal{H}_\Sigma = \mathrm{Ker}(H_S) \subset \hat{\mathcal{H}}_\Sigma \otimes \hat{\mathcal{H}}_S \tag{2.54}$$

Similarly the gluing of two extended Hilbert spaces is obtained by considering the space of singlets under the boundary symmetry, that is by considering the entangling product

$$\mathcal{H}_{\Sigma_L \cup \Sigma_R} = \mathcal{H}_{\Sigma_L} \otimes_{G^S} \mathcal{H}_{\Sigma_R}. \tag{2.55}$$

3 Gravity symplectic potential

We now turn to general relativity. The formalism developed in section 2 generalizes with some additional features. We will see that again the symplectic potential can be made gauge-invariant by addition of a boundary symplectic potential. Just as the boundary symplectic potential in Yang-Mills depends on the choice of gauge at the boundary, the symplectic potential of general relativity depends on a choice of coordinates describing the location of the boundary.

---

[5]When H has a continuous spectrum, the kernel of H may be empty, and one must use the formalism of rigged Hilbert spaces [41] to define the symplectic reduction.

In gravity we have the Lagrangian density (in units where $8\pi G = 1$)

$$L[g] = \tfrac{1}{2}R\epsilon. \tag{3.1}$$

where R is the Ricci scalar of the metric g_{ab}, $\epsilon := \frac{s_g}{D!}\epsilon_{a_1...a_D}\mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_D} = \sqrt{g}\,\mathrm{d}^D x$ is the volume form, and s_g is a sign factor that is -1 when g_{ab} is Lorentzian, and $+1$ when g_{ab} is Euclidean. The symplectic potential current density is given by

$$\theta[g, \delta g] = \frac{1}{2}\nabla_b\left[\delta g^{ab} - g^{ab}\delta g\right]\epsilon_a \tag{3.2}$$

where $\epsilon_a := \frac{s_g}{(D-1)!}\epsilon_{ab_2...b_D}\mathrm{d}x^{b_2} \wedge \cdots \wedge \mathrm{d}x^{b_D} = i_{\partial_a}\epsilon$ is the $(D-1)$-form defining the Hodge duality. Here we used the abbreviated notation $\delta g = g^{ab}\delta g_{ab}$, and $\delta g^{ab} = g^{ac}g^{bd}\delta g_{cd}$ (i.e. δg^{ab} is the variation of the metric with the indices raised, *not* the variation of the inverse metric, which is given by $\delta(g^{ab}) = -\delta g^{ab}$). This follows from the direct calculation of the variation of the Ricci tensor given by

$$\delta R_{ab} = \tfrac{1}{2}\nabla^c\left(\nabla_a\delta g_{bc} + \nabla_b\delta g_{ac} - \nabla_c\delta g_{ab}\right) - \tfrac{1}{2}\nabla_a\nabla_b\delta g \tag{3.3}$$

In the following we will use three notions of equality: $=$ when equality holds off-shell, $\hat{=}$ when the vacuum equation of motion for the background field is satisfied $G_{ab} \hat{=} 0$ but the variation δg_{ab} is arbitrary and $\dot{=}$ when the variation is also restricted to be on shell $\delta G_{ab} \dot{=} 0$.

We now consider how θ transforms under the gauge symmetries of general relativity, the diffeomorphism group. Let $Y : M \to M$ be a diffeomorphism of spacetime and denote by $Y^* : T^*M \to T^*M$ the pullback under this diffeomorphism. We also denote by $Y_* : TM \to TM$ the push forward. Since Y is invertible, the pullback map can be applied to vectors by defining $Y^* = (Y^{-1})_*$ on vectors and then to tensors with both upper and lower indices (see appendix C of [42]). In particular, the pullback of the metric is given by

$$(Y^*g)_{ab}(x) = \partial_a Y^A(x)\partial_b Y^B(x)g_{AB}(Y(x)). \tag{3.4}$$

Just as in Yang-Mills, the variation δ does not commute with the pullback Y^* due to terms that involve the variation of Y. Instead we have the following relation:

$$\delta Y^*(T) = Y^*(\delta T + \mathcal{L}_{\delta_Y}T) \tag{3.5}$$

where $\mathcal{L}_V$ denotes the spacetime Lie derivative along the vector field V and where we have introduced the vector field

$$\delta_Y^a(x) := (\delta Y^a \circ Y^{-1})(x). \tag{3.6}$$

Eq. (3.6) is the generalization of the equations (2.10) and (2.11) of Yang-Mills to the diffeomorphism group. The relation (3.5), which we will use repeatedly, is proved in appendix B for general tensor-valued variational forms.

Now using (3.5) we see that under a diffeomorphism of the underlying fields θ transforms as:

$$\theta[Y^*g, \delta Y^*g] = Y^*\left(\theta[g, \delta g] + \theta[g, \mathcal{L}_{\delta_Y}g]\right). \tag{3.7}$$

This is not simply the pullback of the symplectic potential, because of the second term in (3.7). However we can evaluate it on shell:

$$
\begin{aligned}
\theta[g, \mathcal{L}_{\delta_Y} g] &= \frac{1}{2} \left[\nabla_b \left(\nabla^a \delta_Y^b + \nabla^b \delta_Y^a \right) - 2 \nabla^a \nabla_b \delta_Y^b \right] \epsilon_a \\
&= \left[\left[\nabla_b \nabla^a - \nabla^a \nabla_b \right] \delta_Y^b + \frac{1}{2} \nabla_b \left(\nabla^b \delta_Y^a - \nabla^a \delta_Y^b \right) \right] \epsilon_a \\
&= \left[R^a{}_b \delta_Y^b + \frac{1}{2} \nabla_b (\nabla^b \delta_Y^a - \nabla^a \delta_Y^b) \right] \epsilon_a \\
&\stackrel{.}{=} \frac{1}{2} \mathrm{d} \star \mathrm{d} g \delta_Y .
\end{aligned}
\tag{3.8}
$$

In the last line we evaluate the symplectic potential on-shell, and in the hatted equality $\stackrel{.}{=}$ we have made use of the vacuum Einstein equation $R_{ab} = 0$. For the last term we use the identity

$$
\nabla_b (\nabla^a V^b - \nabla^b V^a) \epsilon_a = \mathrm{d} \star \mathrm{d} g V
\tag{3.9}
$$

where we have used the notation g for the operator of lowering an index on a vector to obtain a 1-form, $gV = V_a \mathrm{d} x^a$, and $\star$ is the Hodge star acting on 2-forms[6] which is given by

$$
\star (\alpha_{ab} \mathrm{d} x^a \wedge \mathrm{d} x^b) = \alpha^{ab} \epsilon_{ab}, \qquad \epsilon_{ab} = \frac{1}{(D-2)!} \epsilon_{abc_1 \ldots c_{D-2}} \mathrm{d} x^{c_1} \wedge \cdots \wedge \mathrm{d} x^{c_{D-2}}.
\tag{3.11}
$$

Here $\epsilon_{ab} = i_{\partial_b} i_{\partial_a} \epsilon$ is the $(D-2)$-form obtained by double contraction with the volume form.

Note that the previous equality can be summarised by introducing a variational vector field $I_V \delta g = \mathcal{L}_V g$. Then the equality becomes

$$
I_V \theta[g, \delta g] = \theta[g, \mathcal{L}_V g] = \mathrm{d} \pi_g[V] + \epsilon_a R^a{}_b V^b \stackrel{.}{=} \mathrm{d} \pi_g[V] .
\tag{3.12}
$$

This is an identity that we will repeatedly use. Here $\pi_g[V]$ is a $(D-2)$-form which depends linearly on V and defined to be

$$
\pi_g[V] := \tfrac{1}{2} \star \mathrm{d} g V .
\tag{3.13}
$$

This shows that under diffeomorphism the symplectic potential current changes by shift under Y and a total derivative:

$$
\theta[Y^* g, \delta Y^* g] = Y^* \left(\theta[g, \delta g] + \mathrm{d} \pi_g[\delta_Y] \right) .
\tag{3.14}
$$

We now consider the symplectic potential Θ_Σ obtained by integrating the current θ over the slice Σ with boundary $\partial \Sigma = S$:

$$
\Theta_\Sigma[g, \delta g] := \int_\Sigma \theta[g, \delta g] .
\tag{3.15}
$$

---

[6]We recall that the convention for Hodge duality on a p-form $\alpha = \frac{1}{p!} \alpha_{a_1 \cdots a_p} \mathrm{d} x^{a_1} \wedge \cdots \wedge \mathrm{d} x^{a_p}$ is

$$
(\alpha \wedge \star \beta) = \frac{1}{p!} \alpha_{a_1 \cdots a_p} \beta^{a_1 \cdots a_p} \epsilon
\tag{3.10}
$$

where ϵ is the volume form.

Θ_Σ is not invariant under diffeomorphism, instead we have

$$\Theta_\Sigma[Y^*g, \delta Y^*g] = \int_{Y(\Sigma)} \theta[g, \delta g] + \int_{Y(S)} \pi_g[\delta_Y]. \tag{3.16}$$

To arrive at this expression we have used the integral identity

$$\int_{Y(\Sigma)} \omega = \int_\Sigma Y^*\omega, \tag{3.17}$$

for a differential form ω. There are two sources of diffeomorphism non-invariance in this expression. The first is that we integrate over the surface $Y(\Sigma)$ with boundary $Y(S)$: this simply reflects that the diffeomorphism moves the coordinate location of the surface over which we integrate. The second source of diffeomorphism non-invariance is the presence of the second term in (3.16), which can be cancelled by addition of a boundary symplectic potential as we will now show.

As in the case of Yang-Mills, we introduce a new variable into the phase space whose transformation law will cancel the boundary term in the symplectic potential. In gauge theory, this extra degree of freedom was a trivialization of the bundle; in a diffeomorphism-invariant theory, the natural analog is a coordinate system X which we view as a mapping $X : U \to M$ where $U \subset \mathbb{R}^D$ is an open set. X is assumed to be invertible on its image. We assume for simplicity that Σ can be covered with one open set and we denote by $\sigma \subset U$ the preimage of Σ, with boundary $\partial\sigma = s$. Under this map we have $\Sigma = X(\sigma)$ and $S = X(s)$. Under a diffeomorphism $Y : M \to M$, the coordinate system X changes to $Y^*(X) = Y^{-1} \circ X$. We can use this coordinate system to define the symplectic potential $\Theta_\Sigma[g, \delta g]$ as the integral over the slice $\Sigma = X(\sigma)$. Under a diffeomorphism its transformation is given by

$$\Theta_{(Y^{-1}\circ X)(\sigma)}[Y^*g, \delta Y^*g] = \Theta_\Sigma[g, \delta g] + \int_S \pi_g[\delta_Y]. \tag{3.18}$$

Thus when integrating over the surface Σ the only change in the symplectic structure under diffeomorphism is via the boundary term.

In order to cancel this boundary term in the transformation law of Θ_Σ, we introduce an additional boundary symplectic potential that depends on δX. The pullback of X by Y does not commute with the variation δ. Instead we find, defining $\delta_X := \delta X \circ X^{-1}$ as in (3.6) and using the chain rule, that under diffeomorphism δ_X transforms as

$$\delta_X \to \delta(Y^{-1} \circ X) \circ (Y^{-1} \circ X)^{-1} = Y^*(\delta_X - \delta_Y). \tag{3.19}$$

Again this equation is directly analogous to the Yang-Mills case (2.16). Using δ_X we can construct a boundary symplectic potential Θ_S to cancel the action of the diffeomorphism on Θ_Σ. Let us define

$$\Theta_S[g, X, \delta X] = \int_S \pi_g[\delta_X], \tag{3.20}$$

where we recall that $S = X(s)$. Under a diffeomorphism it will transform as

$$\Theta_{Y^{-1}(S)}[Y^*g, Y^{-1} \circ X, \delta(Y^{-1} \circ X)] = \int_S \pi_g[\delta_X - \delta_Y]. \tag{3.21}$$

Addition of this boundary term to the symplectic structure will therefore cancel the unwanted term in the transformation of the bulk symplectic potential (3.14). We therefore have the diffeomorphism-invariant symplectic potential given by

$$\Theta[g, \delta g, X, \delta X] = \Theta_\Sigma[g, \delta g] + \Theta_S[g, X, \delta X]$$
$$= \int_\Sigma \theta[g, \delta g] + \oint_S \pi_g[\delta_X]. \qquad (3.22)$$

The X dependence of this expression is implicit in $\Sigma = X(\sigma)$, $S = X(s)$, and in the definition of δ_X. This symplectic potential satisfies by construction the invariance

$$\Theta[Y^*g, \delta Y^*g, (Y^{-1} \circ X), \delta(Y^{-1} \circ X)] = \Theta[g, \delta g, X, \delta X]. \qquad (3.23)$$

We can express the boundary symplectic potential in a more familiar tensorial form using the binormal to the codimension-2 surface S. Let n^a be a timelike unit normal to S, and s^a a spacelike unit normal. We can then form the unit binormal $n_{ab} = (n_a s_b - n_b s_a)$, which is independent of the choice of normals n and s. The densitised unit binormal is $\epsilon_{ab} = \sqrt{q}\, n_{ab}\, \mathrm{d}^{D-2}\sigma$ where σ^α are coordinates on S, and q is the determinant of the induced metric $q_{\alpha\beta}$ on S. The boundary symplectic potential Θ_S can be expressed as

$$\Theta_S[g, X, \delta X] = \frac{1}{2} \oint_S \star \mathrm{d}g\delta_X = \frac{1}{2} \oint_S \epsilon_{ab} \nabla^a \delta_X^b. \qquad (3.24)$$

3.1 Gauge invariance and Hamiltonian generators

We now want to explore the gauge invariance of the extended symplectic structure. This follows closely the formal derivation given in section 2.2 for Yang-Mills. Let V be a vector field and consider the following variation

$$\delta_V g_{ab} = \mathcal{L}_V g_{ab}, \qquad \delta_V X^a = -V^a. \qquad (3.25)$$

This is the infinitesimal version of the transformation $(g, X) \to (Y^*g, Y^{-1} \circ X)$ where $Y = 1 + V + O(V^2)$. The infinitesimal version of diffeomorphism invariance of the symplectic potential shown in (3.23) translates into the identity

$$L_V \Theta = 0. \qquad (3.26)$$

where the equality is valid off-shell. This shows that the Hamiltonian is given by $-I_V(\Theta)$ since

$$I_V \Omega = L_V \Theta - \delta I_V \Theta = -\delta[I_V(\Theta)]. \qquad (3.27)$$

We are therefore left with the evaluation of

$$I_V \Theta = \int_\Sigma \theta[g, \mathcal{L}_V g] - \int_S \pi_g[V] = \int_\Sigma \epsilon_a R^a{}_b V^b \,\hat{=}\, 0. \qquad (3.28)$$

where we have used the identity (3.12). This shows that the generator of gauge transformation vanishes on shell even for gauge transformations with nontrivial support at the boundary.

3.2 Bulk and boundary symplectic forms

Having found Θ we can now evaluate the symplectic form $\Omega = \delta\Theta$, which determines the Poisson brackets between physical observables. There will be two terms in this expression, one coming from Θ_Σ, and the other from Θ_S. Here there is a new subtlety in the derivation that was not present in the Yang-Mills case. Let us first recall that σ is a surface in $\mathbb{R}^D$ with boundary s and that we define $X(\sigma) = \Sigma$ and $X(s) = S$ as their images in spacetime. The new subtlety is that the bulk symplectic potential is an integral over the surface $X(\sigma)$ and this leads to nontrivial variation of the bulk symplectic potential under variations of X. This is in contrast with Yang-Mills where the bulk symplectic potential (2.13) does not depend on the choice of gauge. This dependence on X leads to $\delta_X \delta g$ terms in the variation of Θ_Σ. However we will see that these terms are all total derivatives, and hence can be expressed as boundary integrals. Once this is done the symplectic form consists of a bulk integral containing $\delta g \delta g$ terms, and a boundary integral containing $\delta g \delta_X$ and $\delta_X \delta_X$ terms. This structure is similar to that of Yang-Mills theory.

In order to vary an expression such as Θ_Σ, which is an integral over the surface $\Sigma = X(\sigma)$, we first use the pullback by X to express it as an integral over σ:

$$\int_\Sigma \theta = \int_\sigma X^*(\theta). \tag{3.29}$$

Then we need only vary the integrand. To vary the pullback of a tensor-valued variational form such as θ we make use of the identity (B.1):

$$\delta X^*(T) = X^*(\delta T + \mathcal{L}_{\delta_X} T) \tag{3.30}$$

Recall that $\delta_X^a = \delta X^a \circ X^{-1}$ is a vector field associated with the variation δX.

Using (3.30) to vary the bulk symplectic potential Θ_Σ, we find

$$\delta\Theta_\Sigma = \int_\Sigma \delta\theta + \int_\Sigma \mathcal{L}_{\delta_X} \theta. \tag{3.31}$$

The first integral is the usual bulk symplectic form for gravity [26]. It is obtained as the metric variation of the symplectic potential Θ_Σ and therefore contains only $\delta g \delta g$ terms:

$$\Omega_\Sigma = \int_\Sigma \delta\theta. \tag{3.32}$$

Explicitly, it is given by (see [43]):

$$\Omega_\Sigma = \int_\Sigma \left(\frac{1}{2} \delta g^{cd} \nabla^a \delta g_{cd} + \frac{1}{2} \delta g^{ab} \nabla_b \delta g - \delta g^{cd} \nabla_d \delta g_c{}^a + \frac{1}{2} \delta g \nabla_b \delta g^{ab} - \frac{1}{2} \delta g \nabla^a \delta g \right) \epsilon_a \quad (3.33)$$

Here we use again the abbreviated notation used in (3.2).

The expression can then be evaluated in terms of the metric tangent to the slices and the normal accelerations [28, 44, 45]. If we choose a slice Σ defined by the condition $T = $ constant and equipped with a unit normal one-form $n_a dx^a = -N dT$, where N is the lapse we can, following [28], write the bulk component of the symplectic structure explicitly

in terms of the induced metric h and extrinsic curvature $K_n^{ab} := h^{ac}h^{bd}(\nabla_c n_d)$. If Σ has no boundary it is given by

$$\Omega_\Sigma = \int_\Sigma \delta\Pi^{ab}\delta h_{ab} + \delta\Pi_T \delta T, \tag{3.34}$$

where the momenta conjugate to (h_{ab}, T) are given by

$$\Pi^{ab} := \sqrt{h}\left(K_n^{ab} - K_n h^{ab}\right), \qquad \Pi_T := -\sqrt{h}\left(\Delta N\right), \tag{3.35}$$

where $\Delta = D_a h^{ab} D_b$ is the spacelike Laplacian. The second term in (3.34) says that the Laplacian of the lapse is canonically conjugate to the time variable. This seems unusual to promote the lapse to a kinematical variable. The reason it is usually not considered is because in the ADM approach, one fixes the foliation from the outset and focuses solely on hypersurface-normal deformations. These are only a subset of the deformations — those that preserve the foliation — hence they correspond to deformations for which $\delta T = 0$. Once we do this and ignore the canonical pair $(\Delta N, T)$ the lapse becomes simply a Lagrange multiplier for the Hamiltonian constraint. Note however that the restriction to hypersurface-normal deformations is the reason why the ADM algebra is not equivalent to the algebra of spacetime diffeomorphisms, since in general a diffeomorphism changes the foliation.

The second term in (3.31) is a mixed term that depends on δ_X and δg. To analyse it we use Cartan's formula,

$$\mathcal{L}_{\delta_X}\theta = i_{\delta_X}\mathrm{d}\theta + \mathrm{d}i_{\delta_X}\theta. \tag{3.36}$$

The first term in (3.36) is proportional to $\mathrm{d}\theta = \star\frac{1}{2}(\nabla^a\nabla^b\delta g_{ab} - \nabla^2\delta g)$ which vanishes for on-shell variations by the trace of the linearized vacuum Einstein equation. It is therefore a pure boundary term when δg is an on-shell variation. In general, when we do not restrict to on-shell variations, we note that the term $i_{\delta_X}\mathrm{d}\theta$ in the symplectic structure is an unfamiliar feature of the canonical formalism for gravity. Because δ_X appears only in the interior product, this term depends on the normal component of δ_X and hence arises from changes of the foliation. As such, it is similar to the δT term discussed in the preceding paragraph and disappears if one restricts to hypersurface-normal deformations only. Clearly, this issue of general deformations deserves further investigation, but for now, and because it doesn't play any role in our boundary study, we are content to consider variations for which this term vanishes.

We can now focus on the boundary terms in the symplectic form. The second term in (3.36) is manifestly a total derivative, and so can be written as an integral over S:

$$\int_\Sigma \mathrm{d}i_{\delta_X}\theta = \int_S i_{\delta_X}\theta. \tag{3.37}$$

We also have to vary the boundary symplectic form, which can be accomplished using (3.30) again:

$$\delta\Theta_S = \frac{1}{2}\int_S \delta(\star\mathrm{d}g\delta_X) + \mathcal{L}_{\delta_X}(\star\mathrm{d}g\delta_X). \tag{3.38}$$

We can simplify this expression by expanding the middle term using Cartan's formula and discard the total derivative, since the surface S has no boundary. The result is

$$\delta\Theta_S = \frac{1}{2}\int_S \delta(\star \mathrm{d}g)\delta_X + i_{\delta_X}\mathrm{d}\star \mathrm{d}g\delta_X + \star \mathrm{d}g\delta(\delta_X). \tag{3.39}$$

Combining this variation of Θ_S with the boundary integral from the variation of Θ_Σ we arrive at the boundary symplectic form:

$$\Omega_S = \frac{1}{2}\int_S 2i_{\delta_X}\theta + i_{\delta_X}\mathrm{d}\star \mathrm{d}g\delta_X + \delta(\star \mathrm{d}g)\delta_X + \star \mathrm{d}g\delta(\delta_X). \tag{3.40}$$

The first and third terms in this equation are the mixed $\delta g\delta_X$ terms, while the second and fourth are the $\delta_X\delta_X$ terms. The bulk piece contains only terms of the form $\delta g\delta g$ if we restrict either to on-shell or to hypersurface-normal deformations. Thus we find a structure similar to that of Yang-Mills theory; the boundary terms in the symplectic structure lead to nontrivial Poisson brackets between X and certain components of the metric in the neighbourhood of S. Note that we can write this symplectic structure using the map $\pi_g[V] = \frac{1}{2}\star \mathrm{d}g(V)$ as

$$\Omega_S = \int_S i_{\delta_X}\left(\theta + \mathrm{d}\pi_g[\delta_X]\right) + \delta(\pi_g)[\delta_X] - \tfrac{1}{2}\pi_g[[\delta_X,\delta_X]] \tag{3.41}$$

A significant difference between Yang-Mills and gravity is that the gravity symplectic potential depends on derivatives of δX on the boundary: this shows that there will be a larger group of physical transformations that transform the metric in a first-order neighbourhood of the surface S. Moreover we expect the Poisson brackets between observables to contain derivatives of the generators, unlike in Yang-Mills.

3.3 Surface symmetries

As in gauge theory, we can distinguish between two types of transformations. The transformations considered in section 3.1 are diffeomorphisms, they simply relabel the points of M and hence are pure gauge. This is reflected in the fact that these transformations are null directions in the symplectic form. However, there is a different class of transformations we can consider that transforms the reference frame X^a. These transformations are genuine physical symmetries of the phase space and have nontrivial generators that we will derive in this section. We will refer to them as surface symmetries, because they act at the entangling surface.

Before describing the surface symmetries, we first recall the action of gauge transformations. A diffeomorphism acts on the metric by the pullback, and acts on X on the left:

$$g_{ab} \to Y^*(g)_{ab}, \qquad X \to Y^*(X) = Y^{-1}\circ X. \tag{3.42}$$

If we let $Y = I + V + O(V^2)$ where V is a vector field on M, then $Y^{-1}\circ X = X - V\circ X + O(V^2)$. This means that infinitesimally we have $I_V\delta X = -V\circ X$, and so $I_V\delta_X = -V$. Thus we have the infinitesimal transformation

$$I_V\delta g_{ab} = \mathcal{L}_V g_{ab}, \qquad I_V\delta_X = -V. \tag{3.43}$$

Recall that X is a map from an open set $U \subset \mathbb{R}^D$ into M, and a diffeomorphism Y is a map $M \to M$ that relabels the points. Because of the transformation rule for X, the point $X(x)$ still refers to the same physical point after a diffeomorphism. This transformation is just a relabelling of points in M, and this explains why it is pure gauge.

Instead we will consider a transformation of the reference system $Z : U \to U$. This corresponds to a change of the reference surface s which is mapped onto S, and of the surroundings of that surface. It acts by changing the labelling of the points, but keeps the dynamical fields unchanged:

$$g_{ab} \to g_{ab}, \qquad X \to X \circ Z. \tag{3.44}$$

Note that like in Yang-Mills the symmetry acts on the opposite side of X from the gauge transformation.

To find the infinitesimal version of this transformation, we let $Z = I + w + O(w^2)$, where w is a vector field on U. Then we have in components $(X \circ Z)^a = X^a + \partial_b X^a w^b + O(w^2)$. This defines a vector field Δ_w on phase space where

$$I_{\Delta_w} \delta g_{ab} = 0, \qquad I_{\Delta_w} \delta_X^a = \partial_b X^a (w^b \circ X^{-1}) := W^a. \tag{3.45}$$

Here we use upper case W to denote the pushforward of the vector field w to a vector field on M, i.e. $W = X_*(w)$, where $X_* : TU \to TM$ is the pushforward. The infinitesimal transformation (3.45) is the analog of the transformation generated by the electric field in Yang-Mills (2.37).

Note that because W depends implicitly on X, it has a nontrivial variation δW, which we will need below. This follows from the identity $X^*(W) = w$, and our formula for the variation of the pullback. Varying this identity we find

$$0 = \delta X^*(W) = X^*(\delta W + \mathcal{L}_{\delta_X} W) \tag{3.46}$$

which implies

$$\delta W = -\mathcal{L}_{\delta_X} W = [W, \delta_X]. \tag{3.47}$$

3.4 Classification of surface symmetries

We now have a boundary symplectic structure for general relativity analogous to that of Yang-Mills theory. We have seen that as in Yang-Mills, the extension of the phase space via the introduction of boundary coordinates allows us to construct two types of canonical transformations: the gauge transformations whose Hamiltonian generators vanish on shell and the surface symmetries that act purely on the boundary variables. The latter are associated with a nonvanishing Hamiltonian generator that we will construct in the following section. The key difference between Yang-Mills and general relativity is that the Yang-Mills symmetry generators are spacetime scalars, while the symmetry generators in general relativity are vectors. This leads to a classification of different types of symmetries, depending on how the vector field W meets the surface S. To fix the idea it is convenient to introduce a local coordinate system (x^i, σ^μ) where $(x^i)_{i=0,1}$ denote coordinates normal

to the surface and $(\sigma^\mu)_{\mu=1,\cdots,D-2}$ denote coordinates tangential to the surface which is located at $x^i = $ constant.

Given an infinitesimal symmetry generated by a vector field W, we can divide its restriction to the surface S into its normal part $W_\perp = W^i_\perp \partial_{x^i}$ and tangential part $W_\parallel = W^\mu_\parallel \partial_{\sigma^\mu}$ as $W = W_\parallel + W_\perp$. The important point is that the contraction of the symplectic structure with the generator Δ_w depends not only on the value of W on S but also on the normal derivatives of $W_\perp$, so that even when the normal component vanishes on S, it will have a nontrivial generator. Vector fields whose parallel component vanish $W_\parallel \stackrel{\circ}{=} 0$ and whose normal component and its first derivative in the normal direction also vanishes $W^i_\perp \stackrel{\circ}{=} \partial_j W^i_\perp \stackrel{\circ}{=} 0$, where $\stackrel{\circ}{=}$ denotes equality on the surface S, are in the kernel of the symplectic structure and therefore pure gauge. Thus we only implement as Hamiltonian transformations those vector fields that are nonvanishing in a first-order neighbourhood of S.

We can classify these surface symmetries into three classes:

- *Surface boosts* These are generated by vector fields with $W_\parallel \stackrel{\circ}{=} 0$ and $W_\perp \stackrel{\circ}{=} 0$ but $\partial_i W^j_\perp \stackrel{\circ}{\neq} 0$. They generate position-dependent linear deformations of the normal plane of S. We call them surface boosts because they are a close analog of the boost transformations in the normal plane (x^0, x^1). Like the usual boost, which leaves invariant a codimension-2 plane, the surface boosts leave S invariant. Unlike the usual boost, the surface boosts are not isometries of any particular background. Moreover we will see that these surface boosts generate a larger $\mathrm{SL}(2,\mathbb{R})$ subalgebra of the two-dimensional linear group at each point of S.

- *Surface diffeomorphisms* Vector fields with $W_\perp \stackrel{\circ}{=} 0$ and $W_\parallel \stackrel{\circ}{\neq} 0$ generate infinitesimal diffeomorphisms that map the surface S onto itself.

- *Surface translations* The transformations with $W_\perp \stackrel{\circ}{\neq} 0$ transform the surface S normal to itself.

Here we will focus on the class of surface-preserving transformations, i.e. the transformations such that $W_\perp \stackrel{\circ}{=} 0$. These consist of the surface boosts and surface diffeomorphisms, which do not move the surface S. This is appropriate to the study of an entangling surface which is kept at a fixed location in spacetime, while we allow the angle at which the foliation Σ meets S to change. Previous works have focused on the surface translations (which have been called "would-be pure gauge" degrees of freedom [46]), but have largely overlooked the surface-preserving generators. We leave a full treatment (treating all surface symmetries together) as an interesting area for future work, and comment on some of the technical challenges in the discussion.

3.5 Surface symmetry generators

We now consider the generators H_w of the surface symmetries Δ_w. These are obtained by solving the equation $\delta H_w = I_{\Delta_w}\Omega$. We first find $I_{\Delta_w}\Omega$, which has both a bulk component and a boundary component. The bulk component $I_{\Delta_w}\Omega_\Sigma$ vanishes, because the

transformation Δ_w vanishes when acting on the metric. We are left with the boundary term, which is:

$$
\begin{aligned}
I_{\Delta_w}\Omega &= I_{\Delta_w}\Omega_S \\
&= I_{\Delta_w}\int_S \Big(i_{\delta_X}\left(\theta + \mathrm{d}\pi_g[\delta_X]\right) + (\delta\pi_g)[\delta_X] - \tfrac{1}{2}\pi_g[[\delta_X,\delta_X]]\Big) \\
&= \int_S \Big(i_W\left(\theta + \mathrm{d}\pi_g[\delta_X]\right) - i_{\delta_X}\mathrm{d}\pi_g[W] - (\delta\pi_g)[W] - \pi_g[[W,\delta_X]]\Big)
\end{aligned}
\tag{3.48}
$$

Here we have made use of the identity $I_{\Delta_w}\delta(\delta_X) = -\tfrac{1}{2}I_{\Delta_w}[\delta_X,\delta_X] = [\delta_X,W]$. Here $(\delta\pi_g)$ denotes the variation of the metric terms that appear in π_g without variation of the argument.

We see that this interior product contains two types of terms, one associated with surface-preserving symmetries and one associated with translations. The first one is the translation contribution

$$
\int_S i_W\Big(\theta + \mathrm{d}\pi_g[\delta_X]\Big).
\tag{3.49}
$$

This term vanishes for surface-preserving symmetries. Indeed for surface-preserving symmetries, W is tangent to S, and hence the interior product $i_W\alpha$ vanishes when pulled back to S, for any form α.

From now on we specialize our analysis to the case of surface-preserving symmetries. In this case the previous term vanishes and we are left only with the surface-preserving contribution

$$
I_{\Delta_w}\Omega = -\int_S \Big((\delta\pi_g)[W] + \mathcal{L}_{\delta_X}\pi_g[W] + \pi_g[[W,\delta_X]]\Big).
\tag{3.50}
$$

where we have used the definition of π_g and discarded a total derivative term in order to write the second term as a Lie derivative. Using the identity (3.47) for δW, and the variational identity (3.30), we clearly see that this contraction is Hamiltonian that is $\delta H_w = I_{\Delta_w}\Omega$, where

$$
H_w = -\int_S \pi_g[W] = -\frac{1}{2}\int_S \star \mathrm{d}gW.
\tag{3.51}
$$

H_w is therefore the Hamiltonian generator of the surface-preserving transformations.

We can now easily find the algebra of the surface-preserving transformations. If v is another surface-preserving transformation, then by direct substitution we have

$$
\{H_v, H_w\} = -I_{\Delta_v}\delta H_w = H_{[v,w]}.
\tag{3.52}
$$

Note that the surface-preserving condition is preserved by the Lie bracket: if v and w are surface-preserving then so is $[v,w]$. We also observe that the surface-preserving symmetries do not acquire a central charge at the classical level. We will give a more explicit characterization of this algebra and its Poisson brackets in the following subsection.

We also note that our Hamiltonian (3.51) coincides with the Noether charge of refs. [47, 48]. However the context here is different: W is not assumed to be an isometry of any fixed background, and S need not be a section of a Killing horizon. Instead $W = X_*(w)$ parametrises an infinitesimal change of coordinate frame of the boundary surface.

3.6 Algebra of surface symmetries

In order to describe explicitly the algebra of surface symmetries and its generators, we make a "2+2" decomposition[7] of the metric in a neighbourhood of S, and evaluate the generator H_w of a surface-preserving symmetry w. In terms of the timelike unit normal n and spacelike unit normal s to S, the generator takes the form

$$
\begin{aligned}
H_w &= -\frac{1}{2} \int_S \sqrt{q} (n^a s^b - n^b s^a) \nabla_a W_b \\
&= -\frac{1}{2} \int_S \sqrt{q} \left(\nabla_{\boldsymbol{n}} (W_\perp \cdot \boldsymbol{s}) - \nabla_{\boldsymbol{s}} (W_\perp \cdot \boldsymbol{n}) - W_\parallel^\mu [\boldsymbol{n}, \boldsymbol{s}]^\nu q_{\mu\nu} \right) \\
&= -\frac{1}{2} \int_S \sqrt{q} \left(\epsilon^{AB} \nabla_{\boldsymbol{n}_A} (W_\perp \cdot \boldsymbol{n}_B) - W_\parallel^\mu q_{\mu\nu} [\boldsymbol{n}, \boldsymbol{s}]^\nu \right).
\end{aligned}
\tag{3.53}
$$

We have introduced the notation $\boldsymbol{n}_A = n_A^a \partial_a$ for the two unit normal vectors $\boldsymbol{n}_0 = \boldsymbol{n}$ and $\boldsymbol{n}_1 = \boldsymbol{s}$, and the antisymmetric Levi-Civita symbol ϵ^{AB} with $\epsilon^{01} = 1$. The vector $[\boldsymbol{n}, \boldsymbol{s}]^\nu q_\nu{}^\mu$ is the projection of the Lie bracket of n and s onto the tangent plane of S and it measures the integrability of the normal planes. Indeed the tangent bundle at S can be composed into the vectors tangent to S and their normal $T_S M = TS \oplus (TS)^\perp$. By the Frobenius integrability condition, the normal sub-bundle $(TS)^\perp$ can be described as the tangent bundle of a normal submanifold of spacetime if and only if the Lie-bracket closes $[(TS)^\perp, (TS)^\perp] \subset (TS)^\perp$. In other words, $[\boldsymbol{n}, \boldsymbol{s}]^\nu q_\nu{}^\mu = 0$ is equivalent to the Frobenius integrability condition that the (n, s) plane is tangent to a 2-dimensional submanifold of spacetime.

To evaluate the Hamiltonian explicitly in terms of components of the metric we introduce coordinates (x^i, σ^μ) adapted to the $2 + 2$ decomposition. Let σ^μ be coordinates on S, and let $(x^i)_{i=0,1}$ be two coordinates in the normal directions. In these coordinates we can parametrize the metric as

$$
\mathrm{d}s^2 = h_{ij} \mathrm{d}x^i \mathrm{d}x^j + q_{\mu\nu} \left(\mathrm{d}\sigma^\mu - A_i^\mu \mathrm{d}x^i \right) \left(\mathrm{d}\sigma^\nu - A_j^\nu \mathrm{d}x^j \right).
\tag{3.54}
$$

Here $q_{\mu\nu}$ is the induced metric on S, h_{ij} is a generalized lapse which defines the normal geometry and A_i^μ is a generalized shift. We can view A_i^μ as a normal connection, since under relabelling $\delta\sigma^\mu = \varphi^\mu$ it transforms as a connection $\delta A_i^\mu = \partial_i \varphi^\mu + [A_i, \varphi]^\mu$, where the bracket is the Lie bracket for vectors tangent to S. The relationship between the normals $\boldsymbol{n}^A$ and the normal coordinates x^i is given by a 2×2 matrix with coefficients n_i^a:

$$
\boldsymbol{n}^A = n_i^A \mathrm{d}x^i, \qquad h_{ij} = n_i^A \eta_{AB} n_j^B,
\tag{3.55}
$$

where $\eta_{AB} = \mathrm{diag}(-1, +1)$ is a flat 2D normal metric. This expresses the fact that $\boldsymbol{n}^A$ is normal to the the level sets of constant x^i. It is also necessary to express the normal vector fields $\boldsymbol{n}_A = n_A^a \partial_a$ in terms of the metric coefficients:

$$
\boldsymbol{n}_A = (n^{-1})_A^i (\partial_i + A_i^\mu \partial_\mu).
\tag{3.56}
$$

Indeed we can readily check that $i_{\boldsymbol{n}_A}(\boldsymbol{n}^B) = \delta_A^B$ and that $i_{\boldsymbol{n}_A}(\mathrm{d}\sigma^\mu - A_i^\mu \mathrm{d}x^i) = 0$.

---

[7]Of course, we are in general dimension D, so it is really a $2 + (D - 2)$ decomposition.

We can now further evaluate the Hamiltonian (3.53) in terms of the 2+2 decomposition of the metric

$$
\begin{aligned}
\nabla_{\boldsymbol{n}_A}(W_\perp \cdot \boldsymbol{n}_B) &= (n^{-1})^i_A(\partial_i + A_i^\mu \partial_\mu)(W_\perp^j n_{jB}) \\
&\overset{\circ}{=} (n^{-1})^i_A n_{jB}(\partial_i + A_i^\mu \partial_\mu)(W_\perp^j) \\
&\overset{\circ}{=} h_{jk}(n^{-1})^i_A(n^{-1})^k_B(\partial_i W_\perp^j)
\end{aligned}
$$
(3.57)

where we have used that $W_\perp \overset{\circ}{=} 0$. Contracting with ϵ^{AB} we find

$$
\epsilon^{AB}\nabla_{\boldsymbol{n}_A}(W_\perp \cdot \boldsymbol{n}_B) \overset{\circ}{=} -\frac{1}{\det(n)}h_{jk}\epsilon^{ki}(\partial_i W_\perp^j).
$$
(3.58)

It will be convenient to introduce the densitised normal metric

$$
H_j{}^i := \frac{\sqrt{q}}{\det(n)}h_{jk}\epsilon^{ki}.
$$
(3.59)

It is traceless by construction. Moreover, since $\det(h) = -(\det n)^2$, and $\det(\epsilon) = 1$, we find that

$$
\det H = -\det q, \qquad \mathrm{Tr}(H) = 0.
$$
(3.60)

Using the form (3.56), we can also evaluate the twist density vector

$$
\sqrt{q}[n,s]^\mu q_\mu{}^\nu = \sqrt{q}(n^{-1})^i_0(n^{-1})^j_1[(\partial_i + A_i^\mu\partial_\mu),((\partial_j + A_j^\nu\partial_\nu)]^\nu
$$
(3.61)

$$
= \frac{\sqrt{q}}{\det(n)}\left(\partial_0 A_1^\nu - \partial_1 A_0^\nu + [A_0,A_1]^\nu\right) =: F^\nu.
$$
(3.62)

It is proportional to the curvature of the normal connection. The proportionality coefficient $\sqrt{q}/\det(n)$ is the same we used to rescale the normal metric. From (3.59), (3.62) we see that the Hamiltonian can be written in terms of the densitised normal metric and curvature as

$$
H_w = \frac{1}{2}\int_S \left(H_i{}^j(\partial_j W_\perp^i) + W_\parallel^\mu F_\mu\right).
$$
(3.63)

We can now see that the normal component of the Lie bracket for surface-preserving transformations becomes the $\mathfrak{sl}(2,\mathbb{R})$ algebra while the tangential component reduces to the $(D-2)$-dimensional diffeomorphism group acting on S:

$$
\partial_i[V,W]^j \overset{\circ}{=} \partial_i V_\perp^k \partial_k W_\perp^j - \partial_i W_\perp^k \partial_k V_\perp^j + V_\parallel^\mu \partial_\mu(\partial_i W_\perp^j) - W_\parallel^\mu \partial_\mu(\partial_i V_\perp^j)
$$
(3.64)

$$
[V,W]^\mu \overset{\circ}{=} V_\parallel^\nu \partial_\nu W_\parallel^\mu - W_\parallel^\nu \partial_\nu V_\parallel^\mu.
$$
(3.65)

from which we find the Poisson brackets between generators:

$$
\tfrac{1}{2}\{H_i{}^j(\sigma), H_k{}^l(\sigma')\} = (\delta_i^l H_k{}^j - H_i{}^l \delta_k^j)(\sigma)\delta^{(D-2)}(\sigma,\sigma'),
$$
(3.66)

$$
\tfrac{1}{2}\{F_\mu(\sigma), F_\nu(\sigma')\} = F_\mu(\sigma')\partial_\nu\delta^{(D-2)}(\sigma,\sigma') - F_\nu(\sigma)\partial'_\mu\delta^{(D-2)}(\sigma,\sigma'),
$$
(3.67)

$$
\tfrac{1}{2}\{H_i{}^j(\sigma), F_\mu(\sigma')\} = H_i{}^j(\sigma')\partial_\mu\delta^{(D-2)}(\sigma,\sigma').
$$
(3.68)

We see that the components of the densitised curvature F act as generators of tangential diffeomorphisms. They have a nontrivial action on $H_i{}^j$, which transforms as a scalar under

diffeomorphism, while $H_i{}^j$ itself generates a local $\mathfrak{sl}(2,\mathbb{R})$ algebra. Thus we see that the group of surface-preserving symmetries is $\mathrm{Diff}(S) \ltimes \mathrm{SL}(2,\mathbb{R})^S$.

Let us introduce $\mathfrak{sl}(2,\mathbb{R})$ generators:

$$K_0 = \tfrac{1}{2}(H_0{}^1 - H_1{}^0), \qquad K_1 = H_0{}^0 = -H_1{}^1, \qquad K_2 = \tfrac{1}{2}(H_0{}^1 + H_1{}^0). \tag{3.69}$$

K_0 is an elliptic generator while K_1 and K_2 are hyperbolic. They satisfy the $\mathfrak{sl}(2,\mathbb{R})$ commutation relations

$$\{K_0, K_1\} = 2K_2, \qquad \{K_1, K_2\} = -2K_0, \qquad \{K_2, K_0\} = 2K_1. \tag{3.70}$$

The determinant of H is

$$\det H = H_0^0 H_1^1 - H_0^1 H_1^0 = K_0^2 - K_1^2 - K_2^2 \tag{3.71}$$

which is precisely the Casimir of $\mathfrak{sl}(2,\mathbb{R})$. The unitary representations of $\mathrm{SL}(2,\mathbb{R})$ come in a discrete series for which the Casimir is positive and continuous series for which it is negative. Since we require $\det H$ to be negative, we are interested in the continuous series. Note that the previous relation means that the Casimir of $\mathrm{SL}(2,\mathbb{R})$ is the area element

$$\sqrt{q} = K_1^2 + K_2^2 - K_0^2. \tag{3.72}$$

Since the local $\mathrm{SL}(2,\mathbb{R})$ symmetry is a new feature, we briefly comment on the relation between this symmetry and the more familiar boost symmetry of Carlip and Teitelboim [49]. In that work it was argued that the near horizon symmetry consists of a single boost, and that the quantity canonically conjugate to the boost is the horizon area density $\sqrt{q}$. The difference between the two symmetry groups comes because our $\mathrm{SL}(2,\mathbb{R})$ transformations are not isometries of the normal geometry. The relation between the two symmetries can be understood as a particular gauge-fixing. The normal metric h_{ij} transforms in the adjoint representation of $\mathrm{SL}(2,\mathbb{R})$, and since the metric must be nondegenerate it will spontaneously break this symmetry. It is natural to fix the metric to the flat metric $h_{ij} = \eta_{ij}$, in which case one should consider the little group of transformations that preserves h_{ij}. This gauge corresponds to the conditions $H_0{}^0 = 0$ and $H_0{}^1 = H_1{}^0 = -1$. The little group has a single unbroken generator K_2 which generates boosts. In this gauge, the generator becomes

$$K_2 = \frac{1}{2}(H_0{}^1 - H_1{}^0) = \sqrt{q}, \tag{3.73}$$

the area element of the entangling surface. Thus we recover Carlip and Teitelboim's result as a gauge fixing of ours.

The situation can be summarized with a simple analogy to the rotation group $\mathrm{SO}(3)$. The Casimir J^2 is of course different from the rotation generator J_z. If we fix the axis of rotation then classically J_z and $\sqrt{J^2}$ coincide, but from the point of view of the representation theory they are of course distinct and we should not conflate the two. In particular, they have different spectra at the quantum level and the representation theory is controlled by the Casimir J^2.

4 Quantization and entanglement

So far we have constructed the algebra of boundary symmetries in Yang-Mills theory and in general relativity. At the quantum level, we expect this algebra to be realized in the usual way as an algebra of operators on a Hilbert space. In Yang-Mills theory, this is simply the algebra of surface charges on the entangling surface. In a lattice realization it is the 'electric center' algebra of ref. [15], and it is represented on a Hilbert space spanned by generalized spin network states used in ref. [18].

In gravity, we do not have an explicit construction of the Hilbert space as we do in Yang-Mills. In constructing such a Hilbert space it is extremely useful to have a group of symmetries as a guide. Here we have constructed a large group of surface symmetries whose representation theory can guide the construction of the Hilbert space of a local region. We therefore give a brief sketch of how to construct representations of the algebra of surface symmetries. This appears to be a rather challenging problem whose full solution we are forced to leave to future work.

4.1 Quantization

Having found the algebra satisfied by the surface symmetries in gravity we can look briefly at its quantization, which amounts to studying representations of this algebra. Our surface preserving symmetry group $\text{Diff}(S) \ltimes \text{SL}(2, \mathbb{R})^S$ has the structure of a semidirect product, whose projective representations have been extensively studied by Mackey [50]. This approach is a generalization of the familiar classification of projective representations of the Poincaré group $\text{SO}(1, 3) \ltimes \mathbb{R}^4$. First one starts with irreducible representations of the translation group $\mathbb{R}^4$, which are labelled by momenta, and then defines the little group as the subgroup of the Lorentz group that preserves the given momentum vector.

In our case we have diffeomorphisms playing the role of the Lorentz group and $\text{SL}(2, \mathbb{R})^S$ playing the role of translations. We first fix an irreducible representation of $\text{SL}(2, \mathbb{R})^S$, which amounts to choosing the value of the Casimir operator at each point of S. We will assume that this choice is made smoothly which means that by the relation (3.72) we specify an area form on S. This area form is analogous to the choice of momentum, and the total area (which is diffeomorphism-invariant) is analogous to the mass. Next we consider irreducible representations of the little group, which is the subgroup of diffeomorphisms that preserve the given area form, which we denote $\text{APD}(S)$, where APD stands for Area Preserving Diffeomorphism. The choice of APD representation is analogous to the spin in the classification of Poincaré group representations. Thus the irreducible representations of our algebra are labelled by a total area, and an irreducible representation of $\text{APD}(S)$. Using this information, one can then construct the irreducible representations of the surface symmetry algebra by the method of induced representations.

The algebra of area-preserving diffeomorphisms $\text{APD}(S)$ of a surface S has been studied particularly in the case where $S = S^2$ (which fortunately is also the most physically relevant case, as it would apply to both black hole and cosmological horizons in a four-dimensional spacetime). In particular its structure constants are known [51], and it is known that for

a surface of genus g there are $2g$ central charges. It would therefore be of great interest to understand more completely the representations of APD(S^2).

4.2 Entangling product

Provided we can construct Hilbert spaces associated to local regions, the next question is how to glue together two such Hilbert spaces along a common boundary to make the Hilbert space of a larger region. The diffeomorphism-invariance of general relativity would appear to be an obstacle, since diffeomorphisms naively can move the location at which the two Hilbert spaces are glued. Here we point out that such a gluing construction has effectively been carried out for a large class of rigorously-constructed diffeomorphism-invariant quantum field theories: topological quantum field theories (TQFTs). Although these theories have no local degrees of freedom, they provide useful models of how we can glue subsystems together in a diffeomorphism-invariant theory.

In the axiomatic approach to TQFT, every closed codimension-1 surface Σ is assigned a Hilbert space $Z(\Sigma)$, and each codimension-0 surface M with boundary ∂M is associated with a state in the Hilbert space $Z(\partial M)$ [52]. These satisfy a set of identities that allow one to calculate the partition function of a closed manifold M by cutting along codimension-1 surfaces and gluing them back together. This structure is sufficient to classify two-dimensional TQFTs [53], but in higher dimensions this is not sufficient. This is essentially because codimension-1 surfaces are more complicated objects as the spacetime dimension increases.

One of the axioms of TQFT states that a disjoint union of two surfaces is assigned a tensor product of Hilbert spaces, $Z(\Sigma_1 \sqcup \Sigma_2) = Z(\Sigma_1) \otimes Z(\Sigma_2)$. One can augment this structure by allowing codimension-1 surfaces to be cut along codimension-2 surfaces, the resulting structure is called an extended TQFT [54]. This additional structure is used to classify higher-dimensional extended quantum field theories in terms of n-categories [55].

In general, the structure associated to a codimension-2 surface is a category, and a codimension-1 surface with boundary is a morphism in its boundary category. However one can understand this better in a simple example, such as two-dimensional Yang-Mills theory [56]. Two-dimensional Yang-Mills is not quite a TQFT according to the usual axioms. First, its Hilbert space is infinite-dimensional, and second it is not invariant under the full diffeomorphism group but only under the subgroup of area-preserving diffeomorphisms. However it is close enough to a topological quantum field theory that much of the axiomatic framework carries over to this case [57].

In 2D Yang-Mills, the Hilbert space associated to a closed codimension-1 manifold (a circle) is the space of square-integrable class functions of the gauge group G. These are functions $\psi(u)$ of the holonomy $u = \exp(i \oint A) \in G$ that are invariant under the gauge transformation $u \to g^{-1}ug$, i.e. $\psi(u) = \psi(g^{-1}ug)$. One can also construct the Hilbert space associated to an interval, which is the space of square-integrable functions of the connection integrated along the interval $u = \exp(i \int A) \in G$. In the case of the interval, the states are not restricted to be class functions. The Hilbert space of an interval then carries extra structure: it carries two representations of G corresponding to left and right multiplication

of the group element u acting as $u \to gu$ and $u \to ug^{-1}$ respectively. This endows the Hilbert space of an interval with the structure of a module over the group algebra of G.

The Hilbert spaces of intervals are not glued together with the usual tensor product of Hilbert spaces. Recall that one defines the tensor product over $\mathbb{C}$ as the cartesian product $\mathcal{H}_1 \times \mathcal{H}_2$ modulo the equivalence $(\alpha v, w) \simeq (v, \alpha w)$ where $\alpha \in \mathbb{C}$. In the case of a tensor product over an algebra $\mathcal{A}$, we define $\mathcal{H}_1 \otimes_\mathcal{A} \mathcal{H}_2$ such that $(va) \otimes w = v \otimes (aw)$ for all elements $a \in \mathcal{A}$. When the algebra is the group algebra $\mathcal{A} = \mathbb{C}(G)$ we denote this entangling product $\mathcal{H}_1 \otimes_G \mathcal{H}_2$. As we have seen this entangling group and the corresponding product naturally appear in Yang-Mills and gravity as the surface symmetry group. Interestingly it has been conjectured in [58] that the slice of a spacetime outside a black hole horizon should have precisely this structure. We see here that it appears in fact generically for an entangling surface in a gauge theory and in gravity.

5 Discussion

We have defined a formalism for associating a classical phase space to a bounded region that applies to both Yang-Mills theory and general relativity. We have restricted our analysis to *entangling surfaces*: codimension-2 surfaces which are fixed in spacetime. By including the choice of gauge as a canonical coordinate, we obtain an extended phase space that is invariant under gauge transformations and which carries a representation of an induced boundary symmetry group G_S. In Yang-Mills theory, the boundary degrees of freedom consist of a section of a principal G-bundle and the normal component of the electric field, which act like surface charges on the boundary. In gravity, the boundary degrees of freedom consist of the location of an embedding surface $s \to X^a(s)$ and its normal and tangential geometry, both intrinsic and extrinsic. The remarkable feature of such a construction is that it selects a subset of the boundary observables to be the charges for this emergent boundary symmetry. This is the normal electric field in Yang-Mills, and the normal conformal metric, normal curvature and area density in gravity. While these components of the fields are commuting variables from the point of view of the bulk theory they become noncommutative in the presence of a boundary, forming a Lie subalgebra. We have seen that the appearance of this boundary symmetry group allow us to define at the quantum level the total Hilbert space associated with the gluing of two region as the entangling product. The entangling product is a generalisation of the tensor product where states are restricted to be singlets under the boundary symmetry. We have described a classical version of this entangling product, which takes the form of a symplectic reduction on the product of two phase spaces.

It is well-known in the case of lattice Yang-Mills that the identification of boundary charges as the normal electric field can be understood as a consequence of the discretisation. In a discretisation two facts are clear. First, the discrete electric field generates the action of a nonabelian group and hence must be noncommutative. Second, one must demand continuity of the electric field, but not of the magnetic field across the discrete surface. This is because the electric and magnetic fields are non-commuting variables and the identification $E_L = E_R$ generates a canonical transformation acting on the magnetic fields

B_L and B_R under which admissible magnetic field observables must be invariant. This means that the discretisation requires a splitting of the boundary data into two halves. One half becomes the nonabelian charges and the other half corresponds to the bulk conjugate variables. We can now appreciate why putting gravity on a lattice in a way that respects gauge symmetry is so challenging: it requires an understanding of how to split in half the metric variables along the discrete lattice, and it also requires us to understand what commutation relations these discrete variables should satisfy. What is remarkable in our analysis is that we recover this splitting in two halves of the boundary variables and their commutation relations from an analysis of gauge invariance. This suggest that our analysis should provide new clues into the problem of discretising gravity in a covariant manner. If one views the process of discretisation as a necessary step in order to define a quantum field theory nonperturbatively, this could be relevant for the construction of a theory of quantum gravity.

In fact this reflexion leads to an interesting conjecture: in the continuum we have seen that preserving gauge symmetry in the presence of boundaries implies the existence of a boundary symmetry on entangling surfaces. One can view the vertices of a discrete spacetime as a set of regions bounded by entangling surfaces, which are glued together along links of the lattice. If we take this point of view we can ask what are the necessary conditions on the discrete variables to ensure that the resulting theory manifests diffeomorphism symmetry in the continuum limit? A natural conjecture in view of our analysis is that each vertex of the discretisation needs to carry a representation of the entangling group G_S. In other words, not only does the boundary symmetry appear as a result of gauge invariance, but gauge invariance could be seen to emerge as the result of preserving the local surface symmetry. This is a new and very tantalising possibility that deserves a proper analysis.

Our analysis so far has been restricted to entangling surfaces at a fixed position in the bulk of the spacetime. We have therefore restricted our attention to surface symmetries that preserve the location of the entangling surface. But the local symmetry group also includes surface translations that move the entangling surface. These transformations introduce new complications; in particular we have not yet found a Hamiltonian generator of surface translations analogous to that for surface-preserving transformations (3.51). We have laid the groundwork for such an analysis, but it needs to be completed. Previous work suggests that one will have to impose boundary conditions that fix the location of the surface in order to have a nondegenerate symplectic structure, as found in [33]. It remains to be seen whether such an invariant specification of the boundary region is possible in general and whether it is necessary and sufficient to complete the symplectic structure.

It would be very interesting to generalise our results to the case of more general surfaces and to understand how our results connects to the results obtained when the boundary is at infinity. Much recent work has focused on asymptotic symmetries of Yang-Mills theories [59, 60] and gravity [61, 62]. In particular, [63] showed how the asymptotic symmetries of electrodynamics may arise by augmenting the phase space with edge modes living at the "edge of infinity". However, so far there is no construction of an asymptotically flat phase space for gravity that contains a representation of the BMS group, except in 2+1 dimen-

sions [64]. The methods developed in this paper may be useful to construct a phase space that includes the soft photon and graviton modes on which the action of the asymptotic symmetries is Hamiltonian. We have seen so far that our method constructs a phase space containing a representation of a large nontrivial subgroup of the diffeomorphism group. It would therefore be interesting to impose asymptotically flat boundary conditions within our formalism, and to see what asymptotic symmetry group is obtained.

Another related question concerns the relationship of our analysis with the usual Israel junction conditions. Our analysis shows that we need to identify the normal geometry across the codimension-2 entangling surface but not the tangential geometry (except for the area density), while the Israel junction condition (in the absence of a surface stress tensor) demands the continuity of the tangential metric and extrinsic curvature across a codimension-1 boundary. While there is no immediate contradiction between these results since they apply to radically different boundary conditions (codimension-2 versus codimension-1 boundaries) they create an interesting tension that has to resolve itself when we consider dynamical boundaries. In order to understand this we first need to complete the canonical analysis of surface translations.

In particular, this matching raises the following issue. In electrodynamics, the division of the Hilbert space requires splitting electric field lines into positive and negative surface charges, which must cancel in pairs to create gauge-invariant states. Does this mean that the corresponding matching condition in general relativity requires the introduction of both positive and negative energy?[8] Here we can provide a resolution of this tension. Consider for example a black hole spacetime with a bifurcation surface divided into two regions along this surface. Then the energy is related to the generator of the Killing flow, which we have found is simply a component of the normal metric. These metric components are matched up to a sign coming from the convention of choosing n to be an outward-facing normal vector. Thus there is no contradiction because the generators of surface boosts do not have any fixed sign. On the other hand, the area element does need to be positive in order for the surface S to be spacelike. However we have identified this as the Casimir of $SL(2, \mathbb{R})$, and so the matching condition imposes that it must be equal on both sides. Thus we see that our matching condition gives the expected positivity properties, and there does not seem to be any problem of negative energy. We expect to find no further contradiction when considering the generators of surface translation, but this remains to be explored.

Here we have focused on Yang-Mills theory and general relativity, but our formalism appears to be flexible enough to handle a very general class of field theories. It would be interesting to apply our formalism to more general theories of gravity, for example with higher curvature couplings. As a simplest example, inclusion of a nonzero cosmological constant would allow us to make contact with gravity in asymptotically AdS space. The algebra we derive is reminiscent of the commutation relations of a surface stress tensor, which is an important part of the AdS/CFT correspondence, though we have so far not imposed any boundary conditions. Another interesting case for our formalism is Chern-Simons theory, where the edge states and their relation to entanglement entropy have been understood

---

[8]We thank Steve Giddings for emphasizing this point.

for some time. There we anticipate nontrivial Poisson brackets between φ, which do not appear in pure Yang-Mills theory. This may help to explain the appearance of quantum group symmetry, and may shed new light on the open problem of discretizing Chern-Simons theory, which has so far only been solved for particular lattices and gauge groups [65, 66].

More generally, our work seems to be pointing toward a certain extension of the formalism of quantum field theory. In topological quantum field theory (TQFT) there is a useful notion of an extended TQFT [54]. In extended TQFT, one associates Hilbert spaces not only to closed codimension-1 surfaces, but also to codimension-1 surfaces with boundary. The codimension-2 boundaries of these surfaces are then given some additional structure that encodes how the Hilbert spaces of two regions with boundary may be glued together to make a larger region. This additional structure seems to be precisely what is needed in order to define entanglement entropy, which suggests the utility of considering extended field theory beyond the TQFT context. In addition to formalizing the notion of extended QFT, it also raises the interesting question of whether a quantum field theory can always be extended in this way and whether such an extension is unique. Theories with electric-magnetic duality may provide an interesting example, as they can be viewed as gauge theories in two different ways, suggesting multiple extensions and hence multiple ways of defining the entropy for such theories. The existence of different ways to define entanglement entropy in the algebraic definition of entropy were noted in [15], and there may be an analogous freedom in the extended Hilbert space definition. Another interesting case are certain quantum Hall states admitting distinct edge phases [67], further suggesting different ways of extending the same TQFT.

Ultimately, we would like to apply these classical considerations to the problem of entanglement entropy in quantum gravity. In order to do this, we need a Hilbert space which is the quantization of our classical phase space. This is a hard problem, but it may be easier than the usual quantization problem because we have a large physical symmetry group. Thus an important step in the quantization is to find the irreducible representations of $\text{Diff}(S) \ltimes \text{SL}(2,\mathbb{R})^S$. These irreducible representations play an important role in the entanglement entropy, as we will review below. In particular, the irreducible representations label superselection sectors of the Hilbert space $\mathcal{H}_\Sigma$. Note that the entanglement entropy for compact groups contains a term that depends on the dimension of the irreducible representations of the gauge group. Since $\text{SL}(2,\mathbb{R})$ is noncompact, all irreducible representations are infinite-dimensional; presumably this leads to a divergence that must be regulated. However the regularization of this divergence via heat kernel methods has been understood in other models with $\text{SL}(2,\mathbb{R})$ symmetry [68], so we do not expect this to pose any great difficulty.

Another important question that affects the quantization is the existence of central charges. These central charges appear at the classical level in AdS_3 gravity [69], and underlie the derivation of BTZ black hole entropy [70] and its generalization to higher dimensions [32]. The derivation of (3.52) shows that there is no classical central term in the algebra of surface-preserving transformations. However the central charge in [32] comes from diffeomorphisms that move the surface, so we may expect central terms to appear once we include the surface translations.

Finally, our work can be seen as a first step towards the problem of defining entanglement entropy in nonperturbative quantum gravity. Once we have identified the Hilbert spaces $\mathcal{H}_\Sigma$ carrying an action of the group G_S of surface symmetries, we can define the entanglement entropy by the embedding

$$\mathcal{H} = \mathcal{H}_\Sigma \otimes_{G_S} \mathcal{H}_{\overline{\Sigma}} \subset \mathcal{H}_\Sigma \otimes_{\mathbb{C}} \mathcal{H}_{\overline{\Sigma}}. \tag{5.1}$$

We can now define a density matrix ρ_Σ by embedding states of $\mathcal{H}$ into $\mathcal{H}_\Sigma \otimes \mathcal{H}_{\overline{\Sigma}}$, and tracing over $\mathcal{H}_{\overline{\Sigma}}$. Any density matrix ρ_Σ defined in this way commutes with the action of G_S. This leads to a large symmetry of the problem, and a significant simplification of the entanglement entropy. To make use of this symmetry, we first decompose the Hilbert space into irreducible representations of G_S. Letting R denote the irreducible representations, we have $\mathcal{H}_\Sigma = \bigoplus_R \mathcal{H}_{\Sigma,R}$. Since the reduced density matrix commutes with the action of the surface symmetries, we can also decompose it as $\rho_\Sigma = \bigoplus_R p(R)\rho_{\Sigma,R}$. Here $p(R)$ is the probability that the state lies in the superselection sector described by the representation R; this factor is needed to keep the normalization condition $\mathrm{Tr}(\rho_\Sigma) = 1$ and $\mathrm{Tr}(\rho_{\Sigma,R}) = 1$. The entanglement entropy in gauge theory then takes the form [18]:

$$S_{\mathrm{ent}} = \sum_R p(R)\left(-\log p(R) + \log \dim R + S(\rho_{\Sigma,R})\right). \tag{5.2}$$

Thus the representation theory of the surface symmetry group G_S and the identification of its generators already tells us a lot about the entanglement entropy.

It is natural to compare formula (5.2) for the entanglement entropy with the generalized entropy

$$S_{\mathrm{gen}} = \frac{A}{4G} + S_{\mathrm{out}}. \tag{5.3}$$

In both of these equations the entropy is a sum of an entropy associated with degrees of freedom localized away from the entangling surface (the term S_{out} of the generalized entropy and the term $S(\rho_{\Sigma,R})$ of the entanglement entropy) and the expectation value of an operator (the Bekenstein-Hawking entropy term in the generalized entropy, and the term $\log \dim R$ in the entanglement entropy). The quantity $\log \dim R$ is a gauge-invariant observable that depends on the fields at the entangling surface - for example, in Yang-Mills we can determine it explicitly as a functional of the normal electric field. This suggests the possibility that the Bekenstein-Hawking entropy arises as the part of the entanglement entropy that is determined by the representation theory of the diffeomorphism group. In other words, the Bekenstein-Hawking entropy of a black hole could arise from counting the different ways in which the interior can be glued to its exterior; our boundary symplectic form gives the appropriate measure for this counting. There is of course much to be done to establish or refute this conjecture. In particular, we have to study the representation theory of the surface symmetry group to establish the relationship between $\dim R$ and the geometry of the surface Σ. While the appearance of the area density as the Casimir of local $\mathrm{SL}(2,\mathbb{R})$ symmetry is a suggestive, the real test of quantum gravity is to obtain the area law with the correct finite prefactor of $1/4G$.

Acknowledgments

A Field-space exterior calculus

The purpose of this appendix is to gather some technical results about the calculus of exterior (p, q) forms.

In the algebra of spacetime forms we have the exterior derivative d and interior product i_V from which we define the Lie derivative

$$\mathcal{L}_V = i_V \circ d + d \circ i_V \tag{A.1}$$

In the algebra of field space forms we have an analogous exterior (variational) derivative δ and interior product I, which takes as argument an infinitesimal transformation of the fields. In particular we will be interested in the case where its argument is a spacetime Lie derivative, so that $I_V \delta\phi = \mathcal{L}_V \phi$ where ϕ stands for any of the dynamical fields. In particular, we define:

$$I_V \delta g_{ab} = \nabla_a V_b + \nabla_b V_a, \qquad I_V \delta_X^a = -V^a. \tag{A.2}$$

The field-space interior product with a vector field V satisfies the usual anti-Leibniz rule: when α is a field space p-form and β is a field space q-form, we have

$$I_V(\alpha\beta) = (I_V\alpha)\beta + (-1)^p \alpha(I_V\beta). \tag{A.3}$$

Note, however, that the interior product whose argument is a field-space one-form such as δ_X obeys the Liebniz rule

$$I_{\delta_X}(\alpha\beta) = (I_{\delta_X}\alpha)\beta + \alpha(I_{\delta_X}\beta). \tag{A.4}$$

This rule follows from the antisymmetry property $\alpha\beta = (-1)^{pq}\beta\alpha$ and the fact that I_{δ_X} preserves the degree of forms (whereas I_V decreases it).

One can define also the Lie derivative for field-space forms,

$$L_V = I_V\delta + \delta I_V. \tag{A.5}$$

This satisfies again the usual Liebniz rule

$$L_V(\alpha\beta) = (L_V\alpha)\beta + \alpha L_V(\beta). \tag{A.6}$$

For the Lie derivative along δ_X, we must define the Lie derivative with the opposite sign,

$$L_{\delta_X} = I_{\delta_X}\delta - \delta I_{\delta_X}. \tag{A.7}$$

This leads to the graded Liebniz rule

$$L_{\delta_X}(\alpha\beta) = (L_{\delta_X}\alpha)\beta + (-1)^p \alpha(L_{\delta_X}\beta) \tag{A.8}$$

which reflects the fact that L_{δ_X} increases the form degree by one.

We now show that the field-space Lie derivative is related to the spacetime Lie derivative as:

$$L_V = \mathcal{L}_V + I_{\delta V}. \tag{A.9}$$

We first show that it holds for scalars and the coordinate one-forms $\delta\phi$, and then extend this to arbitrary products. For a field-space 0-form f we have

$$L_V(f) = I_V \delta f = \mathcal{L}_V f = (\mathcal{L}_V + I_{\delta V})f \tag{A.10}$$

and a coordinate one-form, we have

$$L_V(\delta\phi) = \delta I_V \delta\phi = \delta(\mathcal{L}_V \phi) = \mathcal{L}_V \delta\phi + \mathcal{L}_{\delta V}\phi = (\mathcal{L}_V + I_{\delta V})\delta\phi. \tag{A.11}$$

Now suppose that forms α and β both obey the identity (A.9). Then their product does also, which follows from the fact that both L_V and $\mathcal{L}_V + I_{\delta V}$ satisfy the Liebniz rule:

$$\begin{aligned} L_V(\alpha\beta) &= L_V(\alpha)\beta + \alpha L_V(\beta) \\ &= (\mathcal{L}_V\alpha + I_{\delta_V}\alpha)\beta + \alpha(\mathcal{L}_V\beta + I_{\delta_V}\beta) \\ &= (\mathcal{L}_V + I_{\delta V})(\alpha\beta). \end{aligned} \tag{A.12}$$

This completes the proof of (A.9).

For a transformation such as L_{δ_X}, the identity (A.9) is slightly modified. In this case L_{δ_X} is an infinitesimal transformation of the fields that depends linearly on δ_X, so it is a vector-valued differential form. The Lie derivative can be naturally extended to vector-valued forms (see [71]) and we can carry out the same generalization to phase space. Since δ_X is a field-space 1-form, L_{δ_X} increases the degree by one and hence unlike L_V obeys an *anti*-Liebniz rule. This leads to the identity:

$$L_{\delta_X} = \mathcal{L}_{\delta_X} - I_{\delta(\delta_X)}. \tag{A.13}$$

To see this note that on a scalar f we have (as before)

$$L_{\delta_X} f = I_{\delta_X}\delta f - \delta(I_{\delta_X}f) = \mathcal{L}_{\delta_X}f. \tag{A.14}$$

However since $\mathcal{L}_{\delta_X}$ and δ are both antiderivations, we have

$$L_{\delta_X}(\delta\phi) = I_{\delta_X}(\delta\delta\phi) - \delta(I_{\delta_X}\delta\phi) = -\delta(\mathcal{L}_{\delta_X}\phi) = \mathcal{L}_{\delta_X}\delta\phi - \mathcal{L}_{\delta(\delta_X)}\phi = (\mathcal{L}_{\delta_X} - I_{\delta(\delta_X)})\delta\phi. \tag{A.15}$$

Thus (A.13) holds for field-space 0-forms and coordinate 1-forms. We can now extend this to all forms, using the fact that L_{δ_X}, $\mathcal{L}_{\delta_X}$ and $I_{\delta\delta_X}$ are all antiderivations, and hence satisfy a graded Leibniz rule. Thus the relation (A.13) is proved for all (p,q) forms.

B Variation of the pullback

Let $X : M \to M$ be a diffeomorphism of spacetime, T be a spacetime tensor and field-space p-form, and let X^*T denote the pullback under diffeomorphism. The purpose of this appendix is to prove the following variational formula:

$$\delta X^*(T) = X^*(\delta T + \mathcal{L}_{\delta_X} T) \tag{B.1}$$

To prove this formula we first establish it on spacetime tensors, and then on a basis of variational 1-forms. We can then extend to arbitrary tensor-valued variational forms by the product rule, using the fact that both δ and $\mathcal{L}_{\delta_X}$ act as derivations on spacetime tensors, and as antiderivations on variational forms.

To prove (B.1) for spacetime tensors, it suffices to prove it for scalars, vectors and covectors and then to extend to tensors of arbitrary rank by taking products. For a scalar function f:

$$\delta X^* f = \delta f(X) = (\delta f)(X) + \delta X^A \partial_A f(X) = X^*(\delta f + \delta_X^a \partial_a f) = X^*(\delta f + \mathcal{L}_{\delta_X} f) \tag{B.2}$$

where we have used that $\delta_X^A = \delta X^A \circ X^{-1}$. Here we temporarily introduce upper-case letter indices on X^A to distinguish them from the lower-case indices on x^a. For a vector v^a:

$$
\begin{aligned}
\delta X^* v^a &= \delta \left(\frac{\partial x^a}{\partial X^A} v^A(X) \right) \\
&= \frac{\partial x^a}{\partial X^A}(\delta v^A)(X) - \frac{\partial x^a}{\partial X^B}\frac{\partial \delta X^B}{\partial x^c}\frac{\partial x^c}{\partial X^A} v^A(X) + \frac{\partial x^a}{\partial X^A}\delta X^B \partial_B v^A(X) \\
&= X^*(\delta v^a - \partial_b \delta_X^a v^b + \delta_X^b \partial_b v^a) \\
&= X^*(\delta v^a + \mathcal{L}_{\delta_X} v^a)
\end{aligned}
\tag{B.3}
$$

For a covector w_a:

$$
\begin{aligned}
\delta X^* w_a &= \delta(\partial_a X^A w_A(X)) \\
&= \partial_a X^A(\delta w_A)(X) + \partial_a \delta X^A w_A(X) + \partial_a X^A \delta X^B \partial_B w_A(X) \\
&= X^*(\delta w_a + \partial_a \delta_X^b w_b + \delta_X^b \partial_b w_a) \\
&= X^*(\delta w_a + \mathcal{L}_{\delta_X} w_a).
\end{aligned}
\tag{B.4}
$$

This establishes the identity for all spacetime tensors, using the fact that both δ and $\mathcal{L}_{\delta_X}$ are derivations.

To extend the identity to variational forms, it suffices to prove it on a basis of variational 1-forms. Our basis of variational 1-forms will consist of δg_{ab} and δ_X^a. For the metric variation δg_{ab}:

$$
\begin{aligned}
\delta X^* \delta g_{ab} &= \delta(\partial_a X^A \partial_b X^B \delta g_{AB}(X)) \\
&\partial_a \delta X^A \partial_b X^B \delta g_{AB}(X) + \partial_a X^A \partial_b \delta X^B \delta g_{AB}(X) + \partial_a X^A \partial_b X^B \delta X^C \partial_C \delta g_{AB}(X) \\
&= X^*((\partial_a \delta_X^c)\delta g_{cb} + (\partial_b \delta_X^c)\delta g_{ac} + \delta_X^c \partial_c \delta g_{ab}) \\
&= X^*(\mathcal{L}_{\delta_X} \delta g_{ab}).
\end{aligned}
\tag{B.5}
$$

And for δ_X^a:

$$\begin{aligned}
\delta X^*(\delta_X^a) &= \delta\left(\frac{\partial x^a}{\partial X^A}\delta X^A\right) \\
&= -\frac{\partial x^a}{\partial X^B}\frac{\partial \delta X^B}{\partial x^c}\frac{\partial x^c}{\partial X^A}\delta X^A \\
&= X^*(\delta_X^b\partial_b\delta_X^a) \\
&= X^*(\tfrac{1}{2}[\delta_X,\delta_X]^a) \\
&= X^*(\delta(\delta_X^a) + \mathcal{L}_{\delta_X}\delta_X^a)
\end{aligned} \tag{B.6}$$

where we have used $\delta(\delta_X) = -\frac{1}{2}[\delta_X,\delta_X]$ and $\mathcal{L}_{\delta_X}\delta_X = [\delta_X,\delta_X]$.

We can now extend the rule (B.1) to arbitrary tensor valued forms by induction. Suppose that α and β each satisfy (B.1), their product must also:

$$\begin{aligned}
\delta X^*(\alpha\beta) &= \delta(X^*(\alpha)X^*(\beta)) \\
&= \delta X^*(\alpha)X^*(\beta) + (-1)^p X^*(\alpha)\delta X^*(\beta) \\
&= X^*(\delta\alpha + \mathcal{L}_{\delta_X}\alpha)X^*\beta + (-1)^p X^*(\alpha)X^*(\delta\beta + \mathcal{L}_{\delta_X}\beta) \\
&= X^*[(\delta\alpha + \mathcal{L}_{\delta_X}\alpha)\beta + (-1)^p\alpha(\delta\beta + \mathcal{L}_{\delta_X}\beta)] \\
&= X^*[(\delta + \mathcal{L}_{\delta_X})(\alpha\beta)]
\end{aligned} \tag{B.7}$$

This follows from the fact that both $(X^{-1})^* \circ \delta \circ X^*$ and $\delta + \mathcal{L}_{\delta_X}$ are antiderivations, and that an antiderivation is defined by its action on a basis of 1-forms. Thus (B.1) is proved for all tensor-valued variational forms.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](#)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] R.D. Sorkin, *On the Entropy of the Vacuum outside a Horizon*, in *Tenth International Conference on General Relativity and Gravitation, Contributed Papers. Vol. 2*, Padova Italy, 4–9 July (1983), pg. 734 [`arXiv:1402.3589`] [INSPIRE].

[2] L. Bombelli, R.K. Koul, J. Lee and R.D. Sorkin, *Quantum source of entropy for black holes*, *Phys. Rev.* **D 34** (1986) 373 [INSPIRE].

[3] M. Srednicki, *Entropy and area*, *Phys. Rev. Lett.* **71** (1993) 666 [`hep-th/9303048`] [INSPIRE].

[4] S.N. Solodukhin, *Entanglement entropy of black holes*, *Living Rev. Rel.* **14** (2011) 8 [`arXiv:1104.3712`] [INSPIRE].

[5] E. Bianchi and R.C. Myers, *On the Architecture of Spacetime Geometry*, *Class. Quant. Grav.* **31** (2014) 214002 [`arXiv:1212.5183`] [INSPIRE].

[6] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [`arXiv:1005.3035`] [INSPIRE].

[7] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781 [`arXiv:1306.0533`] [INSPIRE].

[8] T. Jacobson, *Thermodynamics of space-time: the Einstein equation of state*, *Phys. Rev. Lett.* **75** (1995) 1260 [gr-qc/9504004] [INSPIRE].

[9] T. Jacobson, *Entanglement Equilibrium and the Einstein Equation*, *Phys. Rev. Lett.* **116** (2016) 201101 [arXiv:1505.04753] [INSPIRE].

[10] S.B. Giddings, D. Marolf and J.B. Hartle, *Observables in effective gravity*, *Phys. Rev.* **D 74** (2006) 064018 [hep-th/0512200] [INSPIRE].

[11] S.B. Giddings, *Hilbert space structure in quantum gravity: an algebraic perspective*, *JHEP* **12** (2015) 099 [arXiv:1503.08207] [INSPIRE].

[12] Z.-C. Gu and X.-G. Wen, *Emergence of helicity ±2 modes (gravitons) from qbit models*, *Nucl. Phys.* **B 863** (2012) 90 [arXiv:0907.1203] [INSPIRE].

[13] D. Harlow, *Wormholes, Emergent Gauge Fields and the Weak Gravity Conjecture*, *JHEP* **01** (2016) 122 [arXiv:1510.07911] [INSPIRE].

[14] R. Haag, *Local quantum physics: fields, particles, algebras*, Texts and Monographs in Physics, Springer-Verlag, Berlin Germany (1992).

[15] H. Casini, M. Huerta and J.A. Rosabal, *Remarks on entanglement entropy for gauge fields*, *Phys. Rev.* **D 89** (2014) 085012 [arXiv:1312.1183] [INSPIRE].

[16] J. Yngvason, *The role of type-III factors in quantum field theory*, *Rept. Math. Phys.* **55** (2005) 135 [math-ph/0411058] [INSPIRE].

[17] P.V. Buividovich and M.I. Polikarpov, *Entanglement entropy in gauge theories and the holographic principle for electric strings*, *Phys. Lett.* **B 670** (2008) 141 [arXiv:0806.3376] [INSPIRE].

[18] W. Donnelly, *Decomposition of entanglement entropy in lattice gauge theory*, *Phys. Rev.* **D 85** (2012) 085004 [arXiv:1109.0036] [INSPIRE].

[19] S. Ghosh, R.M. Soni and S.P. Trivedi, *On The Entanglement Entropy For Gauge Theories*, *JHEP* **09** (2015) 069 [arXiv:1501.02593] [INSPIRE].

[20] S. Aoki, T. Iritani, M. Nozaki, T. Numasawa, N. Shiba and H. Tasaki, *On the definition of entanglement entropy in lattice gauge theories*, *JHEP* **06** (2015) 187 [arXiv:1502.04267] [INSPIRE].

[21] R.M. Soni and S.P. Trivedi, *Aspects of Entanglement Entropy for Gauge Theories*, *JHEP* **01** (2016) 136 [arXiv:1510.07455] [INSPIRE].

[22] W. Donnelly, *Entanglement entropy and nonabelian gauge symmetry*, *Class. Quant. Grav.* **31** (2014) 214003 [arXiv:1406.7304] [INSPIRE].

[23] A. Lewkowycz and J. Maldacena, *Exact results for the entanglement entropy and the energy radiated by a quark*, *JHEP* **05** (2014) 025 [arXiv:1312.5682] [INSPIRE].

[24] W. Donnelly and A.C. Wall, *Entanglement entropy of electromagnetic edge modes*, *Phys. Rev. Lett.* **114** (2015) 111603 [arXiv:1412.1895] [INSPIRE].

[25] W. Donnelly and A.C. Wall, *Geometric entropy and edge modes of the electromagnetic field*, arXiv:1506.05792 [INSPIRE].

[26] C. Crnkovic, *Symplectic Geometry of the Covariant Phase Space, Superstrings and Superspace*, *Class. Quant. Grav.* **5** (1988) 1557 [INSPIRE].

[27] T. Regge and C. Teitelboim, *Role of surface integrals in the hamiltonian formulation of general relativity*, *Annals Phys.* **88** (1974) 286 [INSPIRE].

[28] L. Freidel, *Gravitational Energy, Local Holography and Non-equilibrium Thermodynamics*, *Class. Quant. Grav.* **32** (2015) 055005 [arXiv:1312.1538] [INSPIRE].

[29] A.P. Balachandran, L. Chandar and A. Momen, *Edge states in gravity and black hole physics*, *Nucl. Phys.* **B 461** (1996) 581 [gr-qc/9412019] [INSPIRE].

[30] A.P. Balachandran, L. Chandar and A. Momen, *Edge states in canonical gravity*, gr-qc/9506006 [INSPIRE].

[31] S. Carlip, *The Statistical mechanics of the (2 + 1)-dimensional black hole*, *Phys. Rev.* **D 51** (1995) 632 [gr-qc/9409052] [INSPIRE].

[32] S. Carlip, *Black hole entropy from conformal field theory in any dimension*, *Phys. Rev. Lett.* **82** (1999) 2828 [hep-th/9812013] [INSPIRE].

[33] S. Carlip, *Entropy from conformal field theory at Killing horizons*, *Class. Quant. Grav.* **16** (1999) 3327 [gr-qc/9906126] [INSPIRE].

[34] S. Carlip, *Effective Conformal Descriptions of Black Hole Entropy*, *Entropy* **13** (2011) 1355 [arXiv:1107.2678] [INSPIRE].

[35] J. Kijowski and W. Szczyrba, *A Canonical Structure for Classical Field Theories*, *Commun. Math. Phys.* **46** (1976) 183 [INSPIRE].

[36] C. Crnković and E. Witten, *Covariant description of canonical formalism in geometrical theories*, in *Three hundred years of gravitation*, S.W. Hawking and W. Israel eds., Cambridge University Press, Cambridge U.K. (1987), pg. 676.

[37] K. Gawędzki, *Classical origin of quantum group symmetries in Wess-Zumino-Witten conformal field theory*, *Commun. Math. Phys.* **139** (1991) 201 [INSPIRE].

[38] J. Lee and R.M. Wald, *Local symmetries and constraints*, *J. Math. Phys.* **31** (1990) 725 [INSPIRE].

[39] A. Weinstein, *CBMS Regional Conference Series in Mathematics. Vol. 29: Lectures on Symplectic Manifolds*, American Mathematical Society, Providence U.S.A. (1979).

[40] L. Freidel and A. Perez, *Quantum gravity at the corner*, arXiv:1507.02573 [INSPIRE].

[41] D. Giulini and D. Marolf, *On the generality of refined algebraic quantization*, *Class. Quant. Grav.* **16** (1999) 2479 [gr-qc/9812024] [INSPIRE].

[42] R. Wald, *General Relativity*, University of Chicago Press, Chicago U.S.A. (1984).

[43] S. Hollands and R.M. Wald, *Stability of Black Holes and Black Branes*, *Commun. Math. Phys.* **321** (2013) 629 [arXiv:1201.0463] [INSPIRE].

[44] S.W. Hawking and C.J. Hunter, *Gravitational entropy and global structure*, *Phys. Rev.* **D 59** (1999) 044025 [hep-th/9808085] [INSPIRE].

[45] J.D. Brown, S.R. Lau and J.W. York, Jr., *Action and energy of the gravitational field*, gr-qc/0010024 [INSPIRE].

[46] S. Carlip, *The Statistical mechanics of the (2 + 1)-dimensional black hole*, *Phys. Rev.* **D 51** (1995) 632 [gr-qc/9409052] [INSPIRE].

[47] R.M. Wald, *Black hole entropy is the Noether charge*, *Phys. Rev.* **D 48** (1993) R3427 [gr-qc/9307038] [INSPIRE].

[48] V. Iyer and R.M. Wald, *Some properties of Noether charge and a proposal for dynamical black hole entropy*, *Phys. Rev.* **D 50** (1994) 846 [`gr-qc/9403028`] [ɪɴSPIRE].

[49] S. Carlip and C. Teitelboim, *The off-shell black hole*, *Class. Quant. Grav.* **12** (1995) 1699 [`gr-qc/9312002`] [ɪɴSPIRE].

[50] G. Mackey, *Unitary group representations in physics, probability, and number theory*, *Advanced book classics*, Addison-Wesley, New York U.S.A. (1989).

[51] B. de Wit, J. Hoppe and H. Nicolai, *On the Quantum Mechanics of Supermembranes*, *Nucl. Phys.* **B 305** (1988) 545 [ɪɴSPIRE].

[52] M. Atiyah, *Topological quantum field theories*, *Inst. Hautes Études Sci. Publ. Math.* **68** (1988) 175.

[53] C.J. Schommer-Pries, *The Classification of Two-Dimensional Extended Topological Field Theories*, `arXiv:1112.1000` [ɪɴSPIRE].

[54] J.C. Baez and J. Dolan, *Higher dimensional algebra and topological quantum field theory*, *J. Math. Phys.* **36** (1995) 6073 [`q-alg/9503002`] [ɪɴSPIRE].

[55] J. Lurie, *On the classification of topological field theories*, *Curr. Dev. Math.* **2008** (2009) 129 [`arXiv:0905.0465`].

[56] R. Oeckl, *Two-dimensional quantum Yang-Mills theory with corners*, *J. Phys.* **A 41** (2008) 135401 [`hep-th/0608218`] [ɪɴSPIRE].

[57] S. Cordes, G.W. Moore and S. Ramgoolam, *Lectures on 2 − D Yang-Mills theory, equivariant cohomology and topological field theories*, *Nucl. Phys. Proc. Suppl.* **41** (1995) 184 [`hep-th/9411210`] [ɪɴSPIRE].

[58] E. Witten, *Geometric Langlands From Six Dimensions*, `arXiv:0905.2720` [ɪɴSPIRE].

[59] A. Strominger, *Asymptotic Symmetries of Yang-Mills Theory*, *JHEP* **07** (2014) 151 [`arXiv:1308.0589`] [ɪɴSPIRE].

[60] T. He, P. Mitra, A.P. Porfyriadis and A. Strominger, *New Symmetries of Massless QED*, *JHEP* **10** (2014) 112 [`arXiv:1407.3789`] [ɪɴSPIRE].

[61] A. Strominger, *On BMS Invariance of Gravitational Scattering*, *JHEP* **07** (2014) 152 [`arXiv:1312.2229`] [ɪɴSPIRE].

[62] T. He, V. Lysov, P. Mitra and A. Strominger, *BMS supertranslations and Weinberg's soft graviton theorem*, *JHEP* **05** (2015) 151 [`arXiv:1401.7026`] [ɪɴSPIRE].

[63] A. Mohd, *A note on asymptotic symmetries and soft-photon theorem*, *JHEP* **02** (2015) 060 [`arXiv:1412.5365`] [ɪɴSPIRE].

[64] T. Andrade and D. Marolf, *Asymptotic Symmetries from finite boxes*, *Class. Quant. Grav.* **33** (2016) 015013 [`arXiv:1508.02515`] [ɪɴSPIRE].

[65] V.G. Turaev and O.Y. Viro, *State sum invariants of 3 manifolds and quantum 6j-symbols*, *Topology* **31** (1992) 865 [ɪɴSPIRE].

[66] K. Sun, K. Kumar and E. Fradkin, *Discretized Abelian Chern-Simons gauge theory on arbitrary graphs*, *Phys. Rev.* **B 92** (2015) 115148 [`arXiv:1502.00641`] [ɪɴSPIRE].

[67] J. Cano, M. Cheng, M. Mulligan, C. Nayak, E. Plamadeala and J. Yard, *Bulk-edge correspondence in (2 + 1)-dimensional Abelian topological phases*, *Phys. Rev.* **B 89** (2014) 115116 [`arXiv:1310.5708`] [ɪɴSPIRE].

[68] R.K. Gupta, S. Lal and S. Thakur, *Logarithmic corrections to extremal black hole entropy in* $\mathcal{N} = 2$, *4 and 8 supergravity*, *JHEP* **11** (2014) 072 [arXiv:1402.2441] [iNSPIRE].

[69] J.D. Brown and M. Henneaux, *Central Charges in the Canonical Realization of Asymptotic Symmetries: An Example from Three-Dimensional Gravity*, *Commun. Math. Phys.* **104** (1986) 207 [iNSPIRE].

[70] A. Strominger, *Black hole entropy from near horizon microstates*, *JHEP* **02** (1998) 009 [hep-th/9712251] [iNSPIRE].

[71] I. Kolár, J. Slovák and P.W. Michor, *Natural operations in differential geometry*, Springer-Verlag, Heidelberg Germany (1999).