Poincaré–Plebański formulation of GR and dual simplicity constraints

Vadim Belov

II. Institute for Theoretical Physics, University of Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
E-mail: vadim.belov@desy.de

Received 26 June 2018, revised 24 September 2018
Accepted for publication 19 October 2018
Published 9 November 2018

Abstract
We revise the classical continuum formulation behind the spin foam approach to the quantization of gravity. Based on the recent applications of the current EPRL-FK model beyond triangulations, we identify the tension with the implementation of the ‘volume’ part of simplicity constraints, required for the passage from the topological BF theory to gravity. The crucial role, played by 4D normals in the linear version of constraints, necessitates the extension of the configuration space, and we argue to switch from normal three-forms directly to tetrads. The requirement of vanishing torsion leads to consider first an unconstrained extended Poincaré BF theory, which we characterize fully both at the Lagrangian and Hamiltonian levels, paying special attention to its gauge symmetries. The simplicity constraints are introduced naturally, in the spirit of Plebański formulation, and we give their tetradic version, dual to that of using three-forms. This brings us much closer to the geometric content of general relativity.

Keywords: spin foams, simplicity constraints, topological BF theory, Dirac constrained Hamiltonian analysis, gauge symmetries

(Some figures may appear in colour only in the online journal)
1. Introduction and motivation

Loop quantum gravity (LQG) and spin foams (SF) are the two ambitious non-perturbative approaches to the quantization of gravity. They are usually thought of to be complementary to each other, although not entirely equivalent. The precise correspondence between them is still an open issue. The distinction shows up already at the level of the classical starting points of the two theories, operating with the two dissimilar sets of variables. The first one performs the canonical quantization of the tetradic Holst action [1] (in the time gauge). Whilst the second one is the implementation of the discrete path-integral for gravity [2], based on its Plebański reformulation as a constrained BF theory. The latter naturally generalizes the notion of the state-sum/partition function of topological quantum field theory (TQFT) [3, 4]. It can be used as a way to define the dynamics of LQG, by constructing (the family of) transition amplitudes (for the kinematical states of canonically quantized theory).

In spin foams, the discretized variables of BF theory live on (abstract) two-complexes, typically thought of as being dual to triangulations. The relation to general relativity (GR), in the form of some discrete geometry data, is established through imposition of the simplicity constraints. The current EPRL [5, 6] and FK [7, 8] models are based on the so-called linear formulation, involving the time-normal vector field. In result, the boundary states are spanned by the ‘projected spin-networks’, where these normals, discretized at the nodes, appear as arguments on par with connection variables. The EPRL quantization map thus defines the parameters of the covariant lift of the usual LQG spin-network states [11], and one may regard each SF as a contribution to the sum over histories of canonically quantized geometries [12]. The current work reconsiders the classical formulation of linear simplicity constraints [13], and, in particular, the role of the d.o.f. behind the 4D normals, based on the recent findings in the SF asymptotics.

In the absence of experiment guidance, the major consistency check for the model is the peakedness on the classical geometries, namely on the solutions of the Einstein equations, in a suitable semi-classical regime \( \hbar \to 0 \). Using Barrett’s reconstruction theorem and (extended) stationary phase methods (when the typical scale—the physical area in Planck units—is large), the EPRL-FK amplitude correctly reproduces the four-simplex geometry and Regge gravity, for certain fixed, non-degenerate boundary data [14, 15]. On the other hand, the KKL extension [16, 17].

---

1 The role of these normals has been highlighted by Alexandrov on the grounds of Lorentz covariance, see [9] and references therein; they have also been studied in the context of (extended) group field theory formalism with non-commuting variables [10].
of the simplicial EPRL amplitude to the graphs of arbitrary valence, has not been thoroughly studied until very recently [18], although the partial results were available in the symmetry reduced setting [19]. These results include appearance of certain ‘non-geometric’ configurations, demonstrating shape-mismatch, as well as non-zero physical norm of states with torsion. They are SF analogues of LQG’s ‘twisted’ geometries [20] (discontinuous over flat faces), or torsionful ‘spinning’ (continuous over arbitrarily curved faces) piecewise-flat geometries [21, 22]; whilst Regge configurations appear only as a constrained subset [23].

In section 3.1, we scrutinize the example of hypercuboid and trace back the ‘non-geometricity’ to the way how the simplicity constraints are imposed in the classical theory. In particular, the possibility to neglect the four-volume constraint in the simplex does not hold for more complicated polytopes, so that Barrett’s reconstruction is not applicable. We then proceed with the application to the same system of the fully linear treatment due to Gielen and Oriti [13], with independent normals. The workings of their ‘linear volume’ constraint prompt to switch from the normals (three-forms) directly to edge lengths (tetrads) as new independent variables, using the Hodge duality. In the rest of the paper we study the implications of that change on the classical continuum theory.

First, one may want to incorporate the 4D closure condition on normals (e.g. in the four-simplex [13]), or—as we show to be closely related—the vanishing torsion (i.e. closure for tetrads). In addition to already noted possible link between ‘non-geometricity’ and torsion, it has also been argued on the basis of more involved examples, such as n-point correlation functions and extended triangulations with the bulk curvature. In the latter example, the actual details of taking (semi-)classical limit require to invoke the flipped regime of small Barbero–Immirzi parameter $\gamma \to 0$, in addition to the usual scaling of spins $j \to \infty$. Supposedly, the large $\gamma^{-1}$ in front of the Holst term in the exponent weight of a path integral leads to the dominant contributions from the stationary phase/critical configurations, satisfying the respective equations of motion (e.o.m.)—namely, the simplicial version of the Cartan structure equation: the rotations are trivial (up to $\pi$) in the plane of the hinge. Thus the Immirzi parameter effectively takes control over the strength of another ‘geometricity constraints ... needed in order to reduce the SF dynamical variables to the configurations compatible with the metric geometry of the triangulation’ (see [24] and references therein).

However, neither 4D closure, nor zero-torsion is the part of the theory in the same sense as the bivector closure. Namely, there is no corresponding dynamical law (like the Gauss constraint) from which such condition would follow as (discretized) equations of motion. Therefore we propose in the section 4 the modification of the BF action by the zero-torsion, imposed via Lagrange multipliers. This also seems natural from the point of view of the contact with GR, rather then the Einstein–Cartan (EC) theory. We perform a comprehensive study of this model system prior to imposition of simplicity constraints. This turns out to be topological, with the gauge group being the non-homogeneous Poincaré (affine) extension of the usual homogeneous Lorentz group. The gauge symmetries are defined generically as leaving the action invariant off-shell (up to divergence), and we explicitly derive them from the form of equations of motion in the covariant framework, using the converse of the Noether’s 2nd theorem. The Dirac’s constrained Hamiltonian analysis is performed, having the aim to demonstrate via explicit construction of the gauge generator, which maps solutions of e.o.m.

---

2 The non-zero torsion generically presents in LQG phase space by the lemma 2 in [21], since the Ashtekar–Barbero connection mixes up extrinsic curvature $A = \Gamma[e] + \gamma K$ (its contribution is governed by the Immirzi parameter), residing over the edges of the cellular complex.

3 Such a theory was first considered in [25], as pointed out by the referee. It is closely related (equivalent via integration by parts, i.e. up to boundary terms) to the corresponding topological higher gauge theory, or two-group categorical generalization [26].
onto solutions, that the full 4D symmetry persists on the canonical level as well, although the manifest covariance may be explicitly broken.

At last, the manifest presence of the tetrad frame field $e$ in the formalism among configuration variables—instead of the normal three-forms—makes the introduction of simplicity constraints especially natural. We furthermore propose to look for the linear formulation of [13] in the new guise and introduce its dual version in section 5. The equivalence with the usual simplicity of bivectors is proven. The advantage is in the clear-cut geometric interpretation of the ‘volume’ part of simplicity constraints in terms of an actual three-volume, in analogy to the four-volume of the quadratic case. Besides, the separation between Plebański and 1st order formulations gets blurred to some extent, as their variables are brought together. This resonates nicely with our very first comment on the dissimilarity of the two classical foundations of SF and LQG, respectively.

The structure is as follows. The first half of the paper consists of the brief recap on quadratic Plebański formulation (including the role of four-volume) in section 2, and our revision of the constraints in the EPRL-FK-type models in section 3. Special attention is paid to the role of normals. The familiar reader may skip the exposition and start reading with section 3.1, where the case study of the volume constraints in the EPRL-KKL hypercuboid set-up is performed. This should justify our shift from normals to tetrads in the classical study of the second part of the paper. The section 4 presents the self-contained primer on the Poincaré-BF theory and, as such, can be read independently. The section 5 contains reformulation of linear simplicity in terms of new variables. Finally, we comment on the relations between the various action principles, draw some conclusions in section 6 and discuss on possible outlook for quantization.

2. The recap on constraints: classics

The setting. The classical backdrop behind the spin foam quantization program is the observation due to Plebański [29, 30] that the Einstein–Cartan action can be recast as a constrained BF theory:

$$S[B, \omega, \lambda] = \int_M B_{AB} \wedge F^{AB}[\omega] + \lambda^\alpha C_\alpha[B].$$ (2.1)

Here $A, B = 0, 1, 2, 3$—(internal, or anholonomic) indices in the defining vector representation of the homogeneous Lorentz group $H = SO(3, 1); B$ is the Lie algebra $\mathfrak{h} = \mathfrak{so}(3, 1)$-valued two-form (or $\mathfrak{h} = \mathfrak{so}(4)$ for Euclidean signature spacetime $M$), transforming in the (co)adjoint representation; $\omega$ is the spin-connection with the curvature $F$. The 1st $\int BF$ term, taken on its own, defines a topological field theory without local degrees of freedom; it admits a well-defined exact state sum quantization, discretized over two complexes (à la spin foam).

The 2nd term represents constraints $C_\alpha[B] = 0$ on $B$-field, enforced by the Lagrange multipliers $\lambda^\alpha$ ($\alpha$—multi-index); they effectively reduce the number of independent $B$-components, so that the bivector is given by the (dual) simple product $B = \star e \wedge e$ of (some) tetrad frame field $e$. Hence, on the constraint surface, the theory acquires the form due to Einstein–Cartan:

$$S_{EC}[e, \omega] = \int_M \frac{1}{2} \varepsilon_{ABCD} e^A \wedge e^B \wedge F^{CD}[\omega].$$ (2.2)

In its turn, it is widely accepted as a 1st order formulation of a theory of gravity, since given the equations of motion for $\omega$ in vacuum are satisfied, this renders the theory (on-shell) to the 2nd order tetradic Einstein–Hilbert action: $S_{EC}|_{\delta \omega} = S_{EC}[e, \omega[e]] \equiv S_{EH}[e].$

This strategy was also advocated within the constrained BF CG approach [27], from which our differs in 2 respects: (1) no categorical generalization is implied or required, see [28], (2) exploration of the new form of linear simplicity constraints.
The strategy in the majority of spin foam approaches is to first quantize and then constrain, according to the following route:

1. discretize the classical theory on a piecewise-flat partition of the spacetime $\mathcal{M}$ (most commonly, simplicial);
2. quantize the topological BF part of the discretized theory;
3. impose (a version of) simplicity constraints $C_{\alpha}[B] \approx 0$ directly at the quantum level.

The non-trivial part in constructing SF models for gravity comes from the third step.

The most widely known and well studied is the Plebański’s quadratic set of constraints, existing in two versions:

(a) $B^{AB} \wedge B^{CD} = V \epsilon^{ABCD}$

(b) $\epsilon_{ABCD} B^{AB} B^{CD} = \tilde{V} \epsilon_{abcd} \Rightarrow (2.3)$

They are equivalent, provided the quantity $V$ (resp. $\tilde{V}$) — which is defined by (2.3) through contraction with $\epsilon$ — is non-vanishing [30]. In that case, there are two non-degenerate sectors of solutions:

$$I^\pm : \quad B^{AB} = \pm e^A \wedge e^B, \quad \text{II}^\pm : \quad B^{AB} = \pm \frac{1}{2} \epsilon^{ABCD} e^C \wedge e^D. \quad (2.4)$$

and $V = \frac{1}{2!} \epsilon_{ABCDBAB} f BCD f = \frac{1}{2} V d^4x$ acquires an interpretation of spacetime four-volume. The sectors $II^\pm$ reproduce (2.2) up to the discrete sign ambiguity, while $I^\pm$-sectors give the topological Holst term. The treatment of degenerate case $V = 0$, and relations between sectors can be found in [32].

The discretizations of classically equivalent forms of constraints (2.3) lead to 2, a priori different, SF models. The (a)-case gives the version of the Reisenberger state sum model [33] (corresponding to a self-dual formulation), whereas the case (b) is the most prevailing and leads to the Barrett–Crane (BC) [34, 35] and the new models. The discrete connection is captured by the finite holonomies

$$h_e[\omega] = \exp \left( \int_e \omega^{AB} J_{AB} \right), \quad (2.5)$$

path-ordered along the dual edges $e$. According to the second choice, one associates to the $B$-field the bivectors

$$\int_{S^2} B^{AB}_{f} = \int_{S^2} B^{AB}, \quad (2.6)$$

by integrating over the co-dimension two cells $S_f$ of the piecewise-flat complex, which we label bijectively with the faces $f$ of the dual two-skeleton. Together, they constitute the discretized set of (kinematical) variables of BF theory and Plebański formulation of gravity. In the latter case, the simplicity constraints should be also discretized.

Suppose, our two-complex is dual to a triangulation. Then, depending on the relative position of triangles, the constraints fall into three types:

(i) $\epsilon_{ABCD} B_{f}^{AB} B_{f'}^{CD} = 0$ for each triangle/face $f$—diagonal (or ‘face’) simplicity;

(ii) $\epsilon_{ABCD} B_{f}^{AB} B_{f'}^{CD} = 0$ if two faces share an edge $f \cap f' = e$—cross-simplicity (or ‘tetrahedral’ constraint);

5 Mention that, when the local flatness is not assumed, the integrand should be acted upon by holonomies, referring it to the single source frame, in order to ensure the correct transformation properties of the Lie algebra element under the gauge rotations (see [21]).
(iii) $\epsilon_{ABCD} B^A_{f1} B^D_{f2} =: \tilde{V}_n(f,f')$ for any pair of faces $f,f'$ meeting at the vertex $v$ and spanning four-simplex volume—the so-called volume (or ‘four-simplex’) constraint.

Each of these constraints have different status and are treated accordingly. In particular, they are implemented, respectively, at the level of faces/tetrahedra/four-simplices.

**The closure condition.** In addition, there is usually imposed also the 3D closure

$$\sum_{f \supset e} B^A_{f1} = 0 \quad \forall e. \tag{2.7}$$

This is the consequence of the BF e.o.m. $\nabla [c] B_{ab} = 0$ in the discrete setting: integrate over the 3D volume of the flat tetrahedron $\tau_e$, putting connection to zero via the gauge transform, and use the Stokes’ theorem. It reflects the gauge invariance of the BF theory and gravity. In the canonical picture, this corresponds to the Gauss law constraint (after the symplectic reduction by the cellular flatness constraint [21], restraining local curvature on hinges), and generates the local gauge rotations. Accordingly, in the quantum theory it is usually implemented via group integration, projecting on an invariant subspace.

**The geometric meaning.** The holonomies give the parallel transport of tensors and spinors, taking into account the relative rotation of reference frames between the path endpoints. Regarding the bivector (2.6), when it comes from the metric structure (i.e. the co-tetrad $e$-field, appropriately discretized), then its norm gives the area of the corresponding triangle and the tensor structure encodes the directions of surface, in locally inertial frame of reference. Strictly speaking, such $B$ should not be considered as a variable corresponding to elementary excitations, but rather has a composite nature. The simplicity constraints formulate the necessary and sufficient conditions for the system of bivectors to correspond to the faces of the discrete cell-complex (metric; in our case, of the single four-simplex).

The 1st condition implies that the bivector is simple, i.e. given by the wedge product of two vectors:

$$(i) \quad \Rightarrow \quad B^A_{f1} = E^A_2 E^B_3 \quad \text{or} \quad \star B^A_{f1} = E^A_2 \star E^B_3.$$

If two triangles share a common edge, then the sum of the corresponding two bivectors is also simple:

$$(ii) \quad \Rightarrow \quad B^A_{f1} = E^A_3 E^B_1 \quad \text{or} \quad \star B^A_{f1} = E^A_3 \star E^B_1, \quad \text{and cyclically } \forall f \subset \tau_e.$$

The condition (2.7) states that the geometry of the tetrahedron $\tau_e$, built on vectors $E_1, E_2, E_3$ (or its dual), has a closed boundary. This condition allows a generalization to arbitrary valence and is sufficient to uniquely specify the geometry of a flat polyhedron [36].

An arbitrary set of ten bivectors, satisfying the above three conditions (supplemented with the orientation reversion $B_{AB} = -B_{BA} + \text{some non-degeneracy requirements}$) forms the so-called bivector geometry. The utility of the concept is that it allows to reconstruct the unique flat four-simplex (up to the orientation, translations and inversions), as is shown in [34]. This is the geometrical underpinning behind the construction of the BC model. The role of (iii) is to ensure that the geometries of the tetrahedra fit together to form consistently a 4D geometry, in particular, that the volume of a four-simplex is invariably defined. The volume constraint (iii) is not the part of conditions, defining the bivector geometry, because it is implied by the constraints on the tetrahedral level and the closure. The derivation goes as follows [5, 8]. Label the 5 tetrahedra with $e = 1, \ldots, 5$; the triangle $\triangle_{12}$ is shared by two respective tetrahedra. Using
the closure (2.7), say for tetrahedron 1, and contracting it with all the other bivectors, one can freely swap between triangles, for instance:

$$\epsilon B(\triangle_{12}) \cdot B(\triangle_{45}) + \epsilon B(\triangle_{13}) \cdot B(\triangle_{45}) = -\epsilon B(\triangle_{14}) \cdot B(\triangle_{45}) - \epsilon B(\triangle_{15}) \cdot B(\triangle_{45}) = 0,$$

so that the r.h.s. eliminates on the surface of the simplicity constraints (ii) ⇒ hence (iii) follows.

In the canonical picture parlance, (iii) is interpreted as a ‘secondary’ constraint, which ensures the dynamical conservation of the simplicity constraints (ii). (This is, however, not the statement of the Hamiltonian analysis of the underlying action [37] in the Bergmann’s terminology.) Replacement of (iii) by (2.7) is particularly beneficial for the quantum theory, since the linear in $B$ and local in each tetrahedron closure constraint is much more easier to deal with. By introducing auxiliary normals to tetrahedra, one could incorporate (i) and (ii) into the single ‘linear cross-simplicity’ constraint, retaining the same geometric picture, which led to the new EPRL-FK models [5–8]. We now discuss briefly some details of this construction, as they appear in the literature.

3. On the quantization in new models

There are various ways to arrive at SF partition function (associated with the two-complex $\Upsilon$)

$$Z_{\Upsilon} = \sum_{f} \prod_{f} A_{f} \prod_{e} A_{e},$$

(3.1)

from the classical input laid out above. Roughly they could be captured in two types:

- Relying on the factorization of the representation (3.1), it is sufficient to quantize the geometry of a four-simplex and then to glue such several contributions together. In particular, this route was pursued in the original derivation of the Barrett–Crane (BC) model [34].

  This may be very illuminating in determining the (kinematical) state space of the model. The vertex amplitude determines the graph’s local dynamics and in the canonical picture it would correspond to an expectation value of the Hamiltonian operator on the boundary spin-network state.

  The drawback of the geometric approach is that it is difficult to find the right face and edge amplitudes, and ensure the gluing is consistent.

- A more well-founded complementary approach is based on the discretized path integral, viewed as a sum over (quantum) spacetime histories. The starting point is the BF path-integral measure:

$$Z_{BF} = \int [dB] [d\omega] e^{i \frac{1}{2} \sum_{f} \text{tr} (B_{f} \wedge F_{[\omega]})} = \int [d\omega] \delta (F_{[\omega]}),$$

(3.2)

which is well-defined in our discrete setting as

$$Z_{BF} = \int_{H^{F}\times H^{E}} \prod_{f} dB_{f} \prod_{e} d\omega \exp \left\{ i \sum_{f} \text{tr} (B_{f} \prod_{e \in \partial f} \omega_{e}) \right\} \approx \int_{H^{E}} \prod_{e} d\omega \prod_{f} \delta \left( \prod_{e \in \partial f} \omega_{e} \right).$$

(3.3)

($F$ and $E$ denote the total number of faces and edges, respectively, of the two-complex. The dot over equality sign forewarns that the second delta on the r.h.s. may appear, in general, depending on the actual group $H$ chosen.) Passing from the group elements $\omega_{e}$ to the representation category via the Plancherel theorem, one can recast (3.3) into the (3.1) state-sum form.
Considering the boundary and states on the induced graph $\Gamma = \partial \Upsilon$, one immediately infers, quite generally, that both approaches lead to the kinematical Hilbert spaces spanned by $H$-spin networks for BF theory. The vertex amplitude is then obtained via evaluation of the boundary state on a flat connection. This picture is exact for gravity in 3D, where it is topological (and, thus, discretization independent). However, passing to 4D, the theory should be properly constrained, and this is where the various ambiguities arise.

**Quantizing the bivectors.** The graph $\Gamma$ (cylindrical) state functional depends on connection by virtue of discrete holonomies of $H$. The bivectors act on $H$ as the right/left invariant vector fields, so that using Minkowski spacetime metric $\eta$ we can identify them with the elements of the dual Lie algebra, carrying the natural Poisson structure:

$$\theta : \bigwedge^2 \mathbb{R}^{3,1} \rightarrow \mathfrak{so}(3,1)^*$$

$$E_1 \wedge E_2 \mapsto \theta(E_1 \wedge E_2)(B) := \eta(B \cdot E_1, E_2), \quad B \in \mathfrak{so}(3,1). \quad (3.4)$$

For the following discussion, let us stick to convention, that the bivector $B$ refers, in general, to the symplectic structure—namely, to the kinetic term, involving derivatives $d\omega$ of the connection. It is the overall pre-factor in front of the curvature, and is promoted to tensor operator in certain representation (e.g. using techniques from geometric quantization). We also reserve the label $\Sigma = e \wedge e$ for the face bivector, for which the simplicity constraints are to be formulated, providing any $\Sigma$, satisfying them, with such an interpretation, conversely. This distinction is sensible in the light of the notorious fact that the correspondence between the two quantities is not unique, due to peculiar feature of the Hodge dual $\star$:

$$B_f = \star \Sigma_f + \frac{1}{\gamma} \Sigma_f \mapsto \hat{B}_f \iff \Sigma_f \mapsto \left(\frac{1}{\gamma} \mp \gamma \right)^{-1} \left(\hat{B}_f - \gamma \star \hat{B}_f\right). \quad (3.5)$$

In other words, there are two independent invariant bilinear forms on $\mathfrak{so}(3,1)$, resulting in the Holst action, which is classically equivalent to the vacuum EC theory (on-shell), but may differ quantum mechanically.

The crucial step then is the implementation of a quantum version of the simplicity constraints at the level of state-sum for BF theory (with $\Sigma$ quantized as in (3.5))\(^6\):

$$\hat{C}_\alpha[\Sigma] \approx 0. \quad (3.6)$$

Depending on the first/second class nature of the set (3.6), they should either annihilate the state functionals (à la Dirac), or to be imposed weakly on matrix elements (à la Gupta-Bleuler). This usually leads to restrictions on spin labels $j_f$ and/or intertwiners $\iota_e$ of the boundary Hilbert space states.

**Linear cross-simplicity.** As has been noted, the major ingredient in the new models is the linearization (partial) of the simplicity constraints. It follows directly from the geometric meaning of conditions (i) and (ii), which basically state that four triangles, belonging to the same tetrahedron $\tau$, and described by the area bivectors $\Sigma_{ef}$, lie in one hyperplane.

In the original construction [5–8], one associates a normal discrete 4D vector $V^e_\tau$, $e = 1, ..., 5$, to each of the 5 tetrahedra in the boundary of a four-simplex (assume they are all timelike

\(^6\)Since (2.3) is invariant w.r.t. $\star$ but not $P_\gamma = 1 \pm \frac{1}{\gamma} \star$, it is somewhat perplexing that for finite $\gamma$ both solution sectors (for $\Sigma$) lead to the Holst action for gravity with different effective parameters. This is not really the issue here, since the linear simplicity isolates the sectors in a more efficient way, regardless of the value of $\gamma$.}
\( V_e \in \mathbb{H}_+^2 \cong SL(2, \mathbb{C})/SU(2) \). The quadratic cross-simplicity (ii) is then replaced (rather ad hoc) by the orthogonality requirement on bivectors:

\[
(ii') \quad \forall f \ni e : \quad \Sigma^{AB}_{f \ni e} V_{Re} = 0 \iff I_B (V_e) \triangleright \Sigma_f = 0. \tag{3.7}
\]

The projector on the r.h.s. separates the ‘boost’ components of \( \Sigma \), which are co-aligned with \( V \):

\[
I_B (V)^{AB,CD} := \pm 2 \gamma^{[B} \eta^{D]} C_{V[D]} \quad \text{and} \quad I_B (V)^{AB,CD} := \eta^{[B} \eta^{D]} \equiv 2 \gamma^{[B} \eta^{D]} C_{V[D]}, \tag{3.8}
\]

from the ‘rotational’ part, generating the conjugate \( H_V = h(V) \triangleright SU(2) \) subgroup, which leaves \( V \) invariant. The gain in this new form of constraints is that it excludes the undesired \( I^5 \)-solutions in (2.4), leaving just the mix of the gravitational \( I^5 \)-sectors, as well as degenerate one [38] (which we do not consider here). It thus imposes stronger conditions than quadratic (i) and (ii), which then automatically follow.

We are making two remarks, calling the attention to normals \( V_e \), the prime interest of the current work. The first observation is that the symbolic notation of (3.6) is to be observed with

\[
\tilde{C}_c [\Sigma, V] \approx 0, \tag{3.9}
\]

in order to reflect the introduction of a new geometric objects, in addition to bivectors. The second comment concerns the relation between the two quantities on the constraint surface, namely:

\[
\Sigma^{AB}_f = E_1^{[A} E_2^{B]} = \frac{3}{h_f^2} \ast (N_{f \ni e} \wedge V_e)^{AB} \quad \text{or, conversely} \quad V_e^A = \frac{1}{3} h_f^2 (\star \Sigma_f^{AB}) N_{f \ni e}. \tag{3.10}
\]

Here the edge vectors \( E_{1,2} \perp N, V \) of triangular face \( S_f \) are orthogonal to the (spacelike, \( N^2 = +1 \)) surface normal \( N \), lying within tetrahedron \( \tau_e \). \( N \cdot V = 0 \). In order for \( |V|^2 \equiv V^A V_A \) to be the 3D volume (squared) of \( \tau_e \), the proportionality coefficient is ought to be the height \( h_f^2 = (E_1 \cdot N) \) from the base \( S_f \) to the apex. It appears from the above relations as if neither of \( \Sigma, V \) could be considered more ‘fundamental’, since one can be expressed through the other and vice versa. The resolution of conundrum ultimately lies, not surprisingly, in the composite nature of quantities, both comprised of tetrad d.o.f. This standpoint will pave the way for our extension of field space in section 4, and reformulation of (3.9) in section 5.

**The EPRL map.** Up to this point, the Barbero–Immirzi parameter \( \gamma \) did not partake in the formulation of constraints and should be irrelevant for their geometric content. It, however, plays somewhat mysterious role in quantization and essential for comparison with canonical LQG theory. Let us briefly recap on the basic features of the quantum vertex amplitude which arise from the weak imposition of (the part of) constraints (3.7) on the group \( H \) irreps that live on faces \( f \) of the two-complex \( \mathcal{Y} \), without delving too much into details though:

- The linear cross-simplicity (3.7) is imposed weakly in the gauge-fixed setting, i.e. for the standard normals—either \( V^A_0 = \delta^A_0 \) for spacelike or \( V^A_3 = \delta^A_3 \) for tetrahedra of mixed signature), characterizing the canonical embedding of \( H_0 = SU(2) \) (or \( H_3 = SU(1,1) \)) into \( H \). All the various techniques (such as vanishing matrix elements \( \sim \) master constraint \( \sim \) restriction of coherent state basis to those with the simple expectation values in the semiclassical limit) lead to the relation between 4D and 3D Casimirs:

\[
\frac{c_\gamma^{(2)} (\chi_f)}{2 C_V (j_{\ni e})} \approx \gamma. \tag{3.11}
\]
This defines the embedding map for ‘spins’ \( j \) into decomposition of \( SL(2, \mathbb{C}) \) irreps \( \chi_j \) w.r.t. little group. One nice feature of (3.11) is that its exact implementation [39] projects the spin-connection \( \omega \) in the holonomies (2.5) to the (covariant lift of) Ashtekar–Barbero connection of LQG:

\[
\pi^{(j)} \left( \omega^{AB}_e \mathcal{J}^{(\gamma)}_{AB} \right) \pi^{(j)} = (\gamma) A^I_l \mathcal{J}^{(\gamma)}_{lj}, \quad (\gamma) A^I_l = \frac{1}{2} \varepsilon^{0j}_I \omega_{JK} - \gamma \omega^{0j}_I, \tag{3.12}
\]

where \( \pi^{(j)} \) projects on the \( j \)-irrep of the \( SU(2) \) subgroup, and \( L^I_l = \frac{1}{2} \varepsilon^{0j}_I \mathcal{J}^{(\gamma)}_{JK} \) is the canonical generator of rotations in the corresponding representation.

\( \bullet \) The part of the linear simplicity (3.7) is first class and imposed strongly. Taking into account (3.11), it is equivalent to the (quadratic) diagonal simplicity (i), if the Barbero–Immirzi parameter is included. It relates the \( SL(2, \mathbb{C}) \) Casimirs:

\[
(1 \pm \gamma^2) C^H_{\gamma} (\chi_j) - 2 \gamma C^{(\gamma)}_{\gamma} (\chi_j) \simeq 0, \tag{3.13}
\]

and puts restrictions on allowed ‘simple’ irreps \( \chi_j \). \( \bullet \) The closure condition (2.7) translates into the requirement of the \( H \)-invariance of the amplitude and is ordinarily implemented through the group integration. It obviously encompasses the invariance w.r.t. the little group \( H_0 \) of the embedded \( j \)-states within the tensor product of simple representations, stacked at the tetrahedron \( \tau_e \) bounded by the faces \( \mathcal{S}_f \). Thereby the EPRL embedding map is established:

\[
\Phi_\gamma : \text{Inv}_{SU(2)} \bigotimes_{f \in \mathcal{S}} j_f \rightarrow \text{Inv}_{SL(2, \mathbb{C})} \bigotimes_{f \in \mathcal{S}} \chi_j, \tag{3.14}
\]

where we denoted the representation spaces with their corresponding labellings, for brevity. Thus the boundary state space is labelled by \( SU(2) \) intertwiners glued into spin-networks. The last portion of the integration over the homogeneous space \( H/H_0 \) can be vied as summing over all possible gauge choices for the normals \( \mathcal{V} \in H \triangleright H_0 \), and so restoring the full Lorentz invariance in the gauge-fixed model.

One clearly sees the subsidiary role of \( \mathcal{V} \)'s: in the construction of the model they are treated as ‘unphysical’ gauge choice, which one can specify freely, and later ‘erase’ this information. In effect, \( \mathcal{V} \) allows one to reduce the problem of constraint imposition to the level of little group \( H_0 \) instead of operating directly on the covariant level of the full Lorentz group \( H \). However, let us pinpoint some delicate issues, regarding these normals:

\( \bullet \) For instance, we know that the relative of the time-normal field explicitly appears as non-trivial lapse/shift components in the Lorentz-covariant canonical quantization of the 1st order action (2.2) with the Holst term. It is also an established fact that the boundary states are spanned by the projected spin networks [11]:

\[
\Psi_{\Gamma = \partial \mathcal{T}} \left( \{ h_l \}, \{ \mathcal{V}_n \} \right) = \left\langle \bigotimes_{f \in \mathcal{S}_f} \left( \pi^{(j_f)}_{\gamma_f} D^{(j_f)}_{\mathcal{V}_n} (h_l) \pi^{(j_f)}_{\gamma_f} \right), \bigotimes_{n \in \partial \mathcal{S}_f} \mathcal{V}_n \right\rangle, \tag{3.15}
\]

where this normals play a prominent role and are discretized naturally over the nodes. The state functionals are invariant w.r.t. the covariant Lorentz group action on both sets of variables:

\[
\Psi \left( \{ h_l \}, \{ \mathcal{V}_n \} \right) = \Psi \left( \{ U_{\gamma_f}^{-1} h_f U_{\gamma_f} \}, \{ U_{\gamma_f} \mathcal{V}_n \} \right), \quad \forall U_n \in H. \tag{3.16}
\]
Historically, one of the incentives, which led to FK model [7], was to solve the so-called ‘ultra-locality’ problem with the BC amplitude—namely, the apparent shortage in intertwiner d.o.f., which signified about the limited nature of correlations between neighbouring four-simplices’ geometries. On a more technical level, the resolution of identity, associated with the invariant vector space $X_e := \text{Inv}_H \{ \bigotimes_{f \supset e} \chi_f \}$ at each edge of initial BF spin foam, rewritten in terms of coherent intertwiners (for a moment, $H = \text{Spin}(4) \cong SU(2) \otimes SU(2)$):

$$1_{X_e} = \bigotimes_{\pm} \int d^2 n_e^\pm d j_e^\pm \int dh_{\tau_e}^\pm h_{\tau_e}^\pm |j_f^\pm \cdot n_f^\pm \rangle \langle h_{\tau_e}^\pm |,$$  

(3.17)
is replaced by a projector, where summation is only over those states in the ‘simple’ representations $j^+ = j^-$, which solve the quantum cross-simplicity (3.7) (as expectation values). Specifically, the existence of a common $\hat{u}_e \in SU(2)$ group element is inferred, representing 4D normal $V_e$, which establishes the relation $n^- = -\hat{u}_e \cdot n^+$. The gluing of two four-simplices—via identifying first the geometries, corresponding to their common tetrahedron $\tau_e$, and only then performing an integration—takes into account the missing correlations between neighbouring vertices, sharing an edge. Whereas the unique Barrett–Crane intertwiner is obtained if one integrates separately at each vertex over (then decoupled) geometries. Arguably, the latter identification concerned only an internal 3D geometry of $\tau_e$, encoded in the spins and 3D normals $\{ j, n \}$, corresponding to the (canonically embedded) little group $H_0 = SU(2)$.

The key point of the present work to treat normals as truly independent geometric variable, characterizing the placement of 3D faces in 4D, creates some tension with the implementation of gauge invariance in the EPRL-FK vertex, if the above logic is extended by analogy to $V$. Indeed, the dependence on the subsidiary variable $\hat{u}_e$ is ‘eaten’ by the follow-up $H$-group integration, performed independently at each vertex. The situation is quite similar to the BC intertwiner, so there still may be some d.o.f. left uncorrelated (even though if gauge).

A similar type of arguments have been put forward on the basis of the Lorentz-covariant canonical quantization endeavour [40]. It has been argued that allowing an additional variable $V$ remain unintegrated, the covariant transformation properties (3.16) necessitate a relaxation of the closure condition [41]. The closure of the discrete bivectors is a too restrictive Gauss law, because the gauge transformations should act on the vector variables as well. In the preliminary Hamiltonian analysis of [13], the very same reason led authors to artificially enlarge the phase space by the fictitious momenta, corresponding to $V$. In the section 4 we will see how our modification responds to both these objectives in quite a natural manner.

3.1. The fate of the ‘volume’ constraint

As discussed in section 2, the discretization of (quadratic) volume constraint employs several tetrahedra of the 4-simplex, hence it is usually thought of as consistency condition on time evolution (‘secondary’ constraint). Indeed, (2.8) shows that if the cross-simplicity together with the 3D closure holds true for all tetrahedra, it does not matter which of the face bivectors are used to calculate the volume of the four-simplex. Thus, it is not imposed explicitly in the quantum theory, once the former two are implemented. The same proof using the cable-wire diagrammatic representation of the four-simplex amplitude shows that this indeed holds in the quantum theory as well [2], at least semi-classically.
We notice that the argument heavily relies on the combinatorics of the four-simplex and does not necessarily extend to the generic case of arbitrary two-complex. Explicitly, this appears already in the simple case of (hyper)cuboidal graph \[19\]. There the crudest approximation, using the semiclassical substitute for the exact vertex amplitude, has been studied for the flat (no curvature) rectangular lattice. The expression for the amplitude is a straightforward KKL generalization of the Euclidean EPRL, in the FK representation using coherent states:

\[
\mathcal{A}_\gamma^\pm = \int \prod_e dh_{ee} \ e^{\pm |h_{ee}|} \sim \sqrt{\frac{(2\pi)^{21}}{\det \left( -\partial^2 S(h_{ee}) \right)}} e^{S(h_{ee})}, \quad j \to \infty, \quad (3.18a)
\]

\[
S^{\pm}[h_{ee}] = \mathcal{1} \pm \frac{\gamma}{2} \sum_{ee'} 2j_{(ee')} \ln(-n_{ee'} h_{ee'}^{-1} h_{ee'} |n_{ee}). \quad (3.18b)
\]

Here the summation goes over the (ordered) pairs \((ee') = (f \cap \partial T_e)\)—the (directed) links of a six-valent combinatorial hypercuboidal boundary graph, and the data \(\{j, n\}\) in this symmetry reduced setting was chosen to represent (semi-classically) \(\mathbb{R}^3\)-cuboids:

\[
|\iota\rangle = \int d\alpha \ b^{\frac{1}{2}} \left( \bigotimes_{j=1}^3 [j, n_j] |j', n'_j\rangle, \quad j' \cdot n' = -ji \cdot n, \right. \quad (3.19)
\]

\(\partial^2 S\) denotes the Hessian matrix, evaluated at the critical point \(\partial S(h_{ee}) = \mathcal{R} S(h_{ee}) = 0\).

It turns out that the four-volume of a flat hypercuboid cannot be unambiguously ascribed to a vertex, using the prescription akin to (iii) for four-simplex, where its consistency is guaranteed by (2.8). If the rectangular lattice is geometric (i.e. we are on the solution to simplicity constraints), it is characterized entirely in terms of its edge lengths \(E_i, i = t, x, y, z\), and the unique geometric four-volume can be computed regardless of the faces chosen \(V_i := E_i E_x E_y E_z = \Sigma_{xt} \Sigma_{yt} = \Sigma_{xt} \Sigma_{yt}\), where each area is simply given by the product of the corresponding edge lengths, e.g. \(\Sigma_{xy} = E_x E_y\), etc. Instead, we get six arbitrary areas/spins \(j_{(ij)}\) which do not necessarily satisfy the latter two conditions. Indeed, if we try to proceed like in (2.8), starting with the expression \(j_{xy} j_{xz}\) (depicted by a ‘grasping’ on figure 1) and applying the 3D closure for the spatial cuboid \(\tau_z\), we end up with a tautological result: the contributions from parallel faces (bounding \(\tau_z\) and the two adjacent anti-podal cuboids \(\tau_y, \tau_x - |\) enter with equal areas/spins but opposite signs \(\epsilon \Sigma_{xy} \cdot \Sigma_{yt} = -\epsilon \Sigma_{xz} \cdot \Sigma_{zt}, i, j, k = x, y, z\), thus contracting each other in the sum’, so we arrive at the dull equality \(j_{xy} j_{xz} = j_{xy} j_{xz}\).

The essential ingredient of the EPRL construction, namely, that one could effectively replace the ‘volume’ part of the simplicity by the 3D closure, is not valid for a higher valence. We encounter the problem that the model is not constrained enough to complete the reduction from BF to gravitational theory. The measure of deviation is captured by the ‘non-geometricity’ parameter, in this case:

\[
\varsigma = \frac{(j_{yz} j_{zt} - j_{xz} j_{yt})}{(j_{xy} j_{zt} - j_{xz} j_{yt})}. \quad (3.20)
\]

The numerical studies of \[19\] show that the non-geometric configurations with \(\varsigma \neq 0\) do generically contribute to the path-integral, although their impact might be exponentially suppressed. The dumping is controlled by the width of the Gaussian—the effective ‘mass’ term

\(\text{Stronger, } \epsilon \Sigma_{xz} \cdot \Sigma_{yt} = \epsilon \Sigma_{xt} \cdot \Sigma_{zt} = 0 \text{ by the cross-simplicity.}\)
Figure 1. Diagrammatic representation of the hypercuboidal amplitude. The notation is as in [2]: lines (‘wires’) depict the $SL(2,\mathbb{C})$ representation matrices, and boxes (‘cables’)—invariant projectors (group integrations). Note: apart from the graph’s purely combinatorial properties, the relevance of the crossings for the 4-volume is highlighted in [42].

Figure 2. Elements of linear volume constraint: edge $E_i$ (purple) is shared by the two faces of the spatial cuboid in the middle, bounding it from the cuboids, orthogonal to $x$ (yellow) and $y$ (blue) directions, respectively. Colour scheme corresponds to $\Sigma \cdot \Psi$ pairing.
$m_2^2(\alpha) \approx 2\alpha - 1 > 0$ for $\alpha \gtrsim 0.5$, which depends crucially on the parameter $\alpha$ in the choice of the face amplitude $A_f^{(\alpha)} = ((2^j_f + 1)(2^j'_f + 1))^\alpha$. Reassuringly, in the same range of $\alpha$ indications were given for the tentative continuum limit in the form of a phase transition, with the restoration of the (remnant) diff-invariance. This lead authors to speculate that the freedom in the face amplitude might be restricted on physical grounds, for one should definitely obtain geometric states in the classical limit.

**Fully linear treatment.** Naturally, the two missing constraints to impose in this elucidating example are $\varsigma = 0$, however, it is unevident how to proceed in the most general case. The simplicial ‘four-volume constraint’ makes little sense here, unless appended with some additional requirements (as we tentatively propose in [42], based on the certain type of graph invariants). Being the part of Plebanski’s quadratic formulation, it is also inorganic to the model built on linear constraints. An alternative fully linear formulation was put forward in [13], providing both the continuum version of the cross-simplicity (3.7), as well as the linearized counterpart for the ‘volume’ constraints of the form (3.9). It introduces the basis of three-forms $\vartheta^A$, whose discretization naturally associates 4D normal vectors

$$\vartheta^A_{\tau_e} = \int_{\tau_e} \vartheta^A$$

(3.21)

to tetrahedra $\tau_e$. Although, strictly speaking, neither Plebanski, nor Gielen–Oriti’s linear version were ever formulated beyond triangulations, let us extrapolate the latter to our cuboidal setting, like we did with four-volume.

Analogously to four-simplex, every vertex is the source of four edge vectors, each of which is shared by exactly three cuboids, intersecting along three faces, respectively. Their directions can be identified with that of the eight cuboidal normals $\hat{\mathbf{V}}_i$ (e.g. aligned with $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}'_i = -\hat{\mathbf{x}}_i$ of standard cartesian grid). However, their norms are considered as free parameters (like surface areas $j$). The bivector data is restricted to satisfy

$$(B_{ij} + B_{ij'}) \cdot \hat{\mathbf{V}}_j = 0,$$  

(3.22)

that is opposite faces of 3D cuboid are required to be parallel and of equal area. Indeed, this is precisely the boundary data for coherent intertwiners (3.19), if 3D normals are defined as $(\mathbf{j}\hat{\mathbf{N}})_i = B_{ij} \cdot \hat{\mathbf{V}}_j$. Pick up some vertex, and fix the edge $\hat{\mathbf{V}}_l$, ejecting from it. For every such pair, via straightforward transfer from [13], one writes down the two linearised volume (vector) constraints:

$$\mathbf{V}_i \cdot \Sigma_{il} = \mathbf{V}_j \cdot \Sigma_{jl} = \mathbf{V}_k \cdot \Sigma_{kl} \quad \forall \, i,j,k \neq l.$$  

(3.23)

We have one equality per each 3D cuboid in the cycle, sharing the given edge as a hinge, at which the two consecutive 2D faces intersect. In other words, the volume constraint is discretized at the edge-cuboid pairing (see figure 2). Because of orthogonality, all the terms are proportional to $\mathbf{V}_i$. Let us regroup it at each vertex via formation of sums over edges, orthogonal to one particular direction $\mathbf{V}_k$:

$$\sum_{\{i,j\} \neq k} \mathbf{V}_i \cdot (\star B_{jk}) = 0 \quad \forall \, k.$$  

(3.24)

(Writing $\star B_{ij} = \Sigma_{ij}$, etc, is just the convenient relabelling of face by its two orthogonal directions, corresponding to cuboids adjacent along $S_j$.) Each individual term in the sum is proportional to $\mathbf{V}_{i\neq k}$, and by linear independence in 3D subspace $\perp \mathbf{V}_k$ we can rewrite pre-factors in the form of proportion:

---

8 We exclude the Immirzi parameter from consideration and put $B = \star \Sigma$. From here on all the following discussion is purely classical.
∀ |Vi| |Bk| / |Bj| ∀ i, j ≠ k, \hspace{1cm} (3.25)

the third equality being the consequence of the other two. Taking another such identity for \( l \neq k \) and the same \( i, j \) we arrive at \( |B_k| |B_l| = |B_i| |B_j| \) — the familiar geometricity conditions.

Our proof, however, showed something more than that. The volumes \( |V_i| \) of three cuboids in the cycle around edge \( V_i \) relate as areas of their bases \( |B_k| \). The proportionality coefficient \( |E_k| = |V_i| / |B_k| \) is independent of \( i \) and we infer the existence of heights/edge lengths, invariantly defined in terms of (independent) \( |V_i|, |B_k| \) from any of three cuboids, sharing this edge. The shapes of faces are obviously matching. We essentially exploited self-reciprocal nature of orthogonal lattice, allowing an identification of edges with \( V_i \) up to constants. Had we allowed fluctuations in directions \( V_i \), the two notions would separate, and multiple angles between the lattice and its dual would intervene the formulas.

**Closure of 4D normals.** If we agree to call ‘non-local’ or ‘dynamical’ either the quantities or relations, where objects from several elements of 3D boundary are involved, then the constraint (3.24) is necessarily non-local. Indeed, it collects volumes of three cuboids at their intersection edge, multiplied by the areas that they cut at the beginning of the edge. Gielen and Oriti demonstrated in the case of the four-simplex that these non-trivial equations may be replaced by another non-local condition of 4D closure of normals:

\[ \sum_{e \in n} \lambda^e = 0 \quad \forall \nu, \hspace{1cm} (3.26) \]

together with the usual cross-simplicity (3.7) and 3D closure of bivectors (2.7) at each tetrahedron \( \tau \) (i.e. ‘local’, or ‘kinematical’ constraints).

This result reminds of the *Minkowski’s theorem*, extended to 4D. Recall that the latter asserts the existence of the unique (up to congruence) flat convex polytope, if the set of its face normals \( \{V_i\} \) is given, satisfying (3.26) (see [36] for the 3D case). The non-trivial part is to demonstrate that it is compatible with characterization via simplicity constraints. We did not find the way to prove the analogous statement for hypercuboid. The complication, arising in this case, we relate to the presence of well-separated boundary elements (e.g. ‘past’ and ‘future’ cuboids in the foliation picture). Whereas in four-simplex, every tetrahedron shares at least one common point with any other.

Returning to constrained-BF framework of spin foam models, our inspection highlights the following fact. In order to complete the reduction from the topological theory to gravity, using the linear formulation (3.9) with independent normals, some additional requirements have to be met. This may be either ‘linear volume’ constraint, or its equivalent. The obvious rewriting of the norms \( |V_i| \), using (3.25), as

\[ |E_i| |B_k| = |E_j| |B_l| = |E_k| |B_i| \quad \forall i, j, k \neq l, \hspace{1cm} (3.27) \]

leads to useful (re)interpretation of constraints (3.23). Namely, they indirectly characterize edge lengths through compatibility both with \( B \)’s and \( V \)’s, such that the unique (metric) volume, equal to \( |V_i| \), can be ascribed to each cuboid. After this identification the normals have to satisfy closure (3.26).

We find it convenient to switch to a simpler variables, leading to a more manageable set of relations. Instead of volumes \( V \), we therefore propose to use directly the edge vectors \( E \), as suggested by (3.27). Unlike the original Gielen–Oriti’s (3.23), the new version can be stated locally at the level of 3D polyhedra. We study the continuum formulation behind it in section 5, proving the equivalence with the usual simplicity constraints \( B = ve \land e \).

As another option, one could prefer to impose the closure on 4D normals \( V \), at least in the case of four-simplex. Ordinarily, the 3D closure (2.7) of bivectors can be thought of as the
result of integration of the continuous Gauss law $\nabla \omega \sum_{AB} = 0$ over a three-ball triangulation (using the Stokes’ theorem and the flatness of connection). There is, however, the problem with such field-theoretic interpretation of (3.26), because the corresponding dynamical law $\nabla_\mu \beta^A = 0$ for independent three-forms is absent in the continuum theory. This flaw is fixed in the section 4. In accord with our choice of variables, we first establish the link with the condition of vanishing torsion for tetrads $e^A$, and then implement it via Lagrange multipliers, prior to any imposition of simplicity constraints.

4. The extended BF action with the frame field

The 3D closure condition (2.7) represents the discrete Gauss law of the BF theory. We are looking for a way to accommodate the discrete 4D closure (3.26) in a similar fashion within a continuum theory through the equations of motion. We restrict ourselves here with the modification of the unconstrained topological theory, postponing the discussion of the simplicity constraints to the next section.

Recall that the three-forms appearing in (3.21) are eventually related to tetrads:

$$\gamma^A = \frac{1}{3!} \gamma^A_{\beta \gamma} \, dx^\beta \wedge dx^\gamma \wedge dx^d = \frac{1}{3!} e^A_{\beta \gamma} \epsilon^\beta_{\mu \nu} e^\gamma_{\nu} e^d_{\mu} \, dx^\beta \wedge dx^\gamma \wedge dx^d = \frac{1}{3!} \epsilon^A_{\beta \gamma \delta} \epsilon^\beta_{\mu \nu} \epsilon^\gamma_{\nu} \epsilon^d_{\mu}.$$

By the Hodge duality (see appendix), we may prefer an alternative parametrization and use directly the tetrad field $e^A_a$, without any loss of generality—the number of components is the same as $\gamma^A_{\beta \gamma}$. In order to identify the appropriate kinetic term, suppose that a piecewise-flat cell complex $\bigcup_v \mathcal{T}_v \simeq \mathcal{M}$, approximating the target spacetime manifold, is given. Express the l.h.s. of (3.26) through Stokes’ theorem:

$$\sum_v \gamma^A_v = \int_{\partial \mathcal{T}_v} \gamma^A = \int_{\partial \mathcal{T}_v} d\gamma^A, \quad \text{where } \partial \mathcal{T}_v = \bigcup_{e \in \mathcal{T}_v} \tau_e \quad (4.2)$$

—the boundary of the four-simplex $\mathcal{T}_v$ dual to the vertex $v$ (more generally, any four-polyhedron). One possibility is to introduce it with the Lagrange multipliers (and covariant derivative for the arbitrary smooth manifold):

$$\int b_A \, D \gamma^A = \int \frac{1}{2} \beta_A \epsilon^A_{\beta \gamma \delta} \epsilon^\beta_{\mu \nu} \epsilon^\gamma_{\nu} \epsilon^d_{\mu} \, dx^\beta \wedge dx^\gamma \wedge dx^d = \beta_A \wedge T^A,$$

where

$$\beta_A := \frac{1}{2} \epsilon_{\beta \gamma \delta} \epsilon^\beta_{\mu \nu} \epsilon^\gamma_{\nu} \epsilon^d_{\mu} \equiv \frac{1}{2} \beta^A_{\beta \gamma} \, dx^\beta \wedge dx^\gamma, \quad \beta^A_{\beta \gamma} = \epsilon^A_{\beta \gamma \delta} \epsilon^\beta_{\mu \nu} \epsilon^\gamma_{\nu} \epsilon^d_{\mu} \quad (4.3)$$

—the two-form taking values in $\mathbb{R}^{3,1}$, whilst $b_A$ are scalar functions (zero-forms). The vanishing of torsion $T^A := De^A$ in the first line is sufficient for $D\beta^A = 0$ (4D closure), and we thus further choose to take generic $\beta_A$ as an independent Lagrange multipliers (conjugate to $e^A$), imposing the stronger condition $De^A = 0^9$. There is still a possibility left for exploration if we had chosen independent $b_A$ and weaker equation $D\beta^A = 0^{10}$.

9 It can be noted that an $n$-form with vanishing exterior derivative and appropriate internal indices can be given a geometric interpretation as describing $n$-simplices closing up to form an $(n + 1)$-simplex. $De^A = 0$ may be seen as another instance where the edges of the 2D triangle add up to zero.

10 In this regard, it would be interesting to compare the two constraints and to identify the sets of compatible geometric configurations in each case (especially in the light of allowed ‘conformal shape-mismatch’ in the recent work [18]).
Before studying gravity per se, let us examine first the simplifying theory, which one gets by modifying correspondingly the pure BF action. Thereby, we suggest to consider the following unconstrained theory:

\[ S_0[\mathbf{B}, \beta, \omega, \mathbf{e}] = \int \mathbf{B}_{AB} \wedge F^{AB} + \beta_A \wedge T^A. \] (4.4)

The main assumption made about the frame field is that it is non-degenerate, and the matrix \( e^a_A \) is invertible. In the form of the action we may recognize the special case of the theory [43, 44], designed to obey the Cartan’s (also Bianchi’s) structure equations (which are the kinematical basis for Riemannian geometry). One could anticipate and announce (4.4) as the Poincaré BF theory—this assertion is made precise, as we develop in this section the structure of the theory in full detail.

**4.1. Gauge symmetries: the Lagrangian approach**

The theories in physics are defined by their symmetries, which are usually presupposed. To find them out from the given action principle we follow the route, converse to the renowned Noether’s 2nd theorem. Namely, the gauge symmetries (by which we understand the dependence of the dynamics on an arbitrary function) manifest themselves through the differential identities among equations of motion. Calculate the Euler–Lagrange derivatives (putting the results into convenient language of forms):

\[
\begin{align*}
\frac{\delta S_0}{\delta B_{AB}} &:= \frac{1}{2!} \epsilon_{abcd} \frac{\delta S_0}{\delta B_{AB}} \, dx^c \wedge dx^d = F^{AB}, \\
\frac{\delta S_0}{\delta \beta_A} &:= \frac{1}{2!} \epsilon_{abcd} \frac{\delta S_0}{\delta \beta_A} \, dx^c \wedge dx^d = T^A, \\
\frac{\delta S_0}{\delta \omega_{AB}} &:= \frac{1}{3!} \epsilon_{abcd} \frac{\delta S_0}{\delta \omega_{AB}} \, dx^b \wedge dx^c \wedge dx^d = \mathbf{D}^{AB} - e^{[A} \wedge \beta^{B]}, \\
\frac{\delta S_0}{\delta e^a_A} &:= \frac{1}{3!} \epsilon_{abcd} \frac{\delta S_0}{\delta e^a_A} \, dx^b \wedge dx^c \wedge dx^d = D\beta^A.
\end{align*}
\]

(4.5)

Setting variations to zero (with some fixed boundary conditions), we obtain the field equations, to which every physical motion must satisfy:

\[ \frac{\delta S_0}{\delta B^{AB}}, \frac{\delta S_0}{\delta \beta^A}, \frac{\delta S_0}{\delta \omega_{AB}}, \frac{\delta S_0}{\delta e^a_A} = 0. \] (4.6)

The torsion now manifestly vanishes (as well as the curvature), and one recognizes in the third line (4.5c) the generalized covariant conservation of \( B \), i.e. the lifted 3D closure of bivectors in the discrete (as promised earlier).

It is now straightforward to derive the corresponding differential relations between functional derivatives (4.5):

\[
\begin{align*}
\mathcal{T}^{(1)}_{AB} &:= D_{B^{AB}} S_0 = 0, \\
\mathcal{T}^{(2)}_{A} &:= D_{\beta^A} S_0 - \epsilon^B \wedge S_0 = 0.
\end{align*}
\]

(4.7)
\[ \mathcal{I}^{(3)}_{AB} := D^A \delta S_0 - \delta \omega^{AB} - 2B_{[A} \wedge \delta S_0 \wedge B_{[B]} - \beta_{[A} \wedge \delta S_0 \wedge \beta_{B]} - \epsilon_{[A} \wedge \delta S_0 \wedge \delta S_0_{B]} = 0, \]

\[ \mathcal{I}^{(4)}_A := \frac{\delta S_0}{\delta e^A} - \beta^B \wedge \frac{\delta S_0}{\delta B^{AB}} = 0, \]

which are vanishing identically (off-shell), without using the equations of motion (only their form). To obtain the above relations we used the 2nd and 1st Bianchi’s identities, i.e. the commutation of covariant derivatives:

\[ D^A e = 0, \quad D^A e = 0. \]

From the complete set of independent differential identities (4.7) we now form a generic linear combination, which is identically equal to zero, and integrate by parts:

\[
\int (\Xi^{AB} \wedge \mathcal{I}^{(1)}_{AB} + \xi^A \wedge \mathcal{I}^{(2)}_A + \eta^{AB} \mathcal{I}^{(3)}_{AB} + u^A \mathcal{I}^{(4)}_A) + \int (D \Xi^{AB} - \xi^A \wedge D^B + \eta^{AB} D^B - \xi^A \wedge e^B) + \delta \omega^{AB} - \delta e^B = 0,
\]

for some arbitrary coefficient functions \((\eta^{AB}, u^A)\) and one-forms \((\Xi^{AB}, \xi^A, \xi^B)\). The transformations that leave the action invariant, up to divergence, are readily seen:

\[ \delta \omega^{AB} = -d \eta^{AB} + f^{AB}_{CD,E} \omega^{CD} \xi^{EF}, \]

\[ \delta \xi^A = -d u^A + f^A_{CD,B} (\omega^{CD} u^B - \xi^{EF}) , \]

\[ \delta B^{AB} = -f^{AB}_{CD,E} \eta^{CD} \beta^E - u^{[A} \beta_{B]} + D \Xi^{AB} - \xi^{AB} \wedge e^B, \]

\[ \delta \beta^A = -f^A_{CD,B} \xi^{CD} \beta^B + D \xi^A, \]

where \( f^{AB}_{CD,E} := \eta_{E[C} \delta^{A]D}_F - \eta_{F[C} \delta^{A]D}_E, \quad f^A_{CD,B} := \eta_{B[C} \delta^A_D). \)

We may combine the connection \( \omega \in \text{so}(3, 1) \) and the gauge potential of translations \( e \) into a single (Cartan) connection \([45]\) of the Poincaré gauge group\(^{11}\):

\[ \text{so}(3, 1) \ltimes \mathbf{p}^{3,1} \ni \varpi := \omega + e = \omega^A J_{AB} + e^C \mathcal{P}_C, \]

whose generators \( J_{AB}, \mathcal{P}_C \) satisfy the algebra (of which (4.8e) are structure constants):

\[ [J_{AB}, J_{CD}] = i \left( \eta_{[A} J_{B]D} - \eta_{D[A} J_{B]C} \right), \]

\[ [J_{AB}, \mathcal{P}_C] = i \eta_{[A} \mathcal{P}_{B]}, \]

\[ [\mathcal{P}_A, \mathcal{P}_B] = 0. \]

A local gauge transformation, taking values in \( G = SO(3, 1) \ltimes \mathbf{p}^{3,1} \), can be split into

\[ g(x) = u(x) U(x), \quad u = e^{-i \varpi}, \quad U = e^{-i \varpi}, \]

\(^{11}\)This is quite appealing from another point of view that the usual non-degeneracy condition on \( \det e^A_a \neq 0 \), which is unclear how to achieve in practise, is also the requirement for the \( g \)-connection to be a Cartan connection.
s.t. \( u(x) \) changes the zero section, i.e. changes the local identification of points of tangency at each spacetime event. The transformation law for the connection is then

\[
\varpi \to \varpi' = g^{-1}(\varpi + d)g = U^{-1}u^{-1}(\omega + e)uU + U^{-1}u^{-1}(du)U + U^{-1}du,
\]

(4.12)

whose infinitesimal form is (4.8a) and (4.8b). The combined curvature transforms in the adjoint representation:

\[
\mathcal{F}[\varpi] \to \mathcal{F}' = g^{-1}\mathcal{F}g = U^{-1}u^{-1}(F + T)uU,
\]

(4.13a)

or, infinitesimally:

\[
\delta F^{AB} = f^{AB}_{CD,EF}F^{CD}q^{EF}, \quad \delta T^A = f^{A}_{CD,B}(F^{CD}u^B - q^{CD}T^B).
\]

(4.13b)

However, the analogous quantity, comprised of the conjugate variables

\[
\mathfrak{so}(3,1) \ltimes \mathfrak{p}^{3,1} \ni B := B + \beta = B^{AB}J_{AB} + \beta^C P_C,
\]

(4.14)

demonstrates slightly different behaviour (4.8c) and (4.8d), compensating for (4.13) in order to make the action invariant. The inhomogeneity is transferred from \( \beta \) to \( B \) part. For instance, if \( \beta \) is of the form (4.3), the corresponding addition would be

\[
-\delta B \ni B^{AB} = \frac{1}{2}(u \cdot b)\mathcal{e}^A \wedge \mathcal{e}^B + \delta A B \wedge (u \cdot e). \quad \text{We are tempted to interpret the latter as being responsible for mixing up the simple bivectors—the symmetry which is usually broken in order to obtain general relativity. In addition to the usual (internal) gauge transformations, the action (4.4) is also invariant w.r.t. the shifts}
\]

\[
\delta \varpi = 0, \quad \delta B = (D\Xi^{AB} - e^A \wedge \xi^B)J_{AB} + D\xi^A P_A \equiv D\varpi(\Xi + \xi),
\]

(4.15)

which extends the usual ‘topological’ BF symmetry. In fact, it is always possible to gauge away any local d.o.f.; however, if \( \mathcal{M} \) is topologically non-trivial, \( \varpi \) and \( \mathcal{B} \) can have non-trivial solutions globally (hence the name).

Despite the non-conventional form of (4.8c), we still have obtained the right connection (4.9) and the algebra (4.10). So that allows us to conclude that we have constructed a topological theory of the BF type for the Poincaré gauge group. We may guess that the departure of \( \mathcal{B} \) transformation properties from the adjoint (4.13) is a forced decision, due to the degenerate nature of the Killing form: for the semidirect product algebra with the abelian ideal of translations it reduces to \( \text{Tr}[\text{ad}_{(U,J)} \circ \text{ad}_{(V,J)}] = 2\text{Tr}[\text{ad}_U \circ \text{ad}_V] \propto \eta_{[A][C][B][D]} = -\frac{1}{4} \xi_{[A][B]} \xi_{[C][D]} \xi_{[E][F]} \xi_{[G][H]}. \)

Thus, the temptation to read the expression in (4.4) as the Cartan–Killing pairing, similar to the \( f(\mathcal{B} \wedge \mathcal{F}) \) with semisimple group, faces obstacles. The contraction for \( \beta = T \) is performed with the \( \mathcal{E} \)-bundle metric, e.g. obtained from the \( \mathcal{J} = \mathcal{P} \) vector couplings. We find this to be an interesting arena for the imposition of simplicity constraints.

**On the diffeomorphisms.** Note that the topological shifts (4.15), parametrized by Lie-algebra valued one-forms, leave the spacetime indices intact, as another manifestation of the topological character of the theory. On the contrary, the diffeomorphisms of \( \mathcal{M} \) are generated by the (horizontal) vector fields \( \xi \in \Gamma(T\mathcal{M}) \), its infinitesimal action on tensors being given by the Lie derivative. Hence, the question of relation between the two symmetries is a subtle one. To clarify the issue one would have to give a definite answer to the following questions: Whether the diffeomorphism transformations constitute an independent symmetry? Since the action functional is construction in a coordinate-free way, one expects it to be invariant under diffeomorphisms. However, among the derived Noether identities we do not immediately find the respective one with the right tensorial structure, which one could contract with the parameter vector field.
One typically finds in the literature [37], the following reply to our query. Of course, we are free to form various linear combinations of (4.7), allowing for the field-dependent coefficients. In such a manner, the following identity can be constructed from the basic ones:

\[
\int \left( \mathbf{B}^{[AB} \mathbf{\omega }_{\mathbf{A]}^B} \right) \wedge I_1 + \left( \mathbf{\beta }^A \right) \wedge I_2 + \left( \mathbf{\omega }_{\mathbf{A]}^B} \mathbf{\omega }_{\mathbf{A]}^B} \right) I_3 + \left( e^A \right) \wedge I_4 = 0, \tag{4.16}
\]

where we contract with the vector field \( \mathbf{\zeta } = \mathbf{\zeta }^a \partial _a \) using the interior product (A.5). This leads to the transformations which are seemingly related to diffeomorphisms:

\[
\begin{align*}
\mathcal{L} \mathbf{\zeta } \mathbf{\omega }_{\mathbf{A]}^B} & = - \delta \mathbf{\omega }_{\mathbf{A]}^B} + \left( \mathbf{\zeta } \frac{\delta S_0}{\delta \mathbf{\omega }_{\mathbf{A]}^B} } \right), \\
\mathcal{L} \mathbf{\zeta } e^A & = - \delta e^A + \left( \mathbf{\zeta } \frac{\delta S_0}{\delta e^A} \right), \\
\mathcal{L} \mathbf{\zeta } B_{AB} & = - \delta B_{AB} + \left( \mathbf{\zeta } \frac{\delta S_0}{\delta \mathbf{\omega }_{\mathbf{A]}^B} } \right), \\
\mathcal{L} \mathbf{\zeta } \mathbf{\beta }^A & = - \delta \mathbf{\beta }^A + \left( \mathbf{\zeta } \frac{\delta S_0}{\delta e^A} \right),
\end{align*}
\tag{4.17}
\]

with gauge parameters as in (4.16); however, the equivalence between the two being valid only on-shell.

This is peculiar, since it tightens the transformation properties to the fact, whether the equations of motion are satisfied. We stress, in this regard, that the diffeos of the conventional second order metric GR are the elementary gauge symmetries in the above sense, both at the Lagrangian and Hamiltonian levels [46]; the corresponding differential identities being that of the 2nd Bianchi’s (contracted). One finds the resolution of conundrum just described somewhat unsatisfactory and will return to this in a separate publication.

4.2. Hamiltonian analysis of the Poincaré BF theory

Having characterized the system completely at the covariant Lagrangian level, we now pass to studying it using the canonical approach. Ordinarily the Hamiltonian methods imply the manifest breaking of covariance by explicitly separating spatial from temporal field-components, and considering equal-time Poisson brackets. We would like to stress that the chosen preferred status of time coordinate in Hamiltonian analysis is not associated \textit{a priori} with an explicit separation of spacetime itself into ‘space and time’ (not at this stage at least), as suggested e.g. by the ADM change of coordinates and their geometrical interpretation. We therefore are being cautious with usage of such notions which are usually referred to as 3+1-decomposition, or slicing/splitting/foliation/etc. In particular, all 4D symmetries persist at the canonical level in the form of gauge generators, mapping solutions into solutions, as we show for this particular example.

Following the Dirac’s general treatment of singular Lagrangian systems [47], one starts by defining the conjugate momenta (for \textit{all} configuration variables):

\[
\begin{align*}
\Pi ^a _{AB} & := \frac{\delta L_0}{\delta \dot{\mathbf{\omega }}_{\mathbf{A]}^B} } \equiv \left( \Pi ^a _{AB}, \Pi ^a _{AB} \right) \approx \left( 0, \frac{1}{2} e^{ijk} B_{ijk} \right), \tag{4.18a} \\
\pi _a^A & := \frac{\delta L_0}{\delta \dot{e }^A } \equiv \left( \pi _a^A, \pi _a^A \right) \approx \left( 0, \frac{1}{2} e^{ijk} \beta ^{ijk} \right). \tag{4.18b}
\end{align*}
\]


where dot denotes the time derivative of the generalized variables \( \dot{q} = \partial_t q \), and \( \epsilon^{ijk} = \epsilon^{ijk}_0 \), \( i,j,k = 1,2,3 \). We have the totality of primary constraints for the generalized coordinates \( q \) and momenta \( p \), in the sense that none of the velocities \( \dot{q} \) enter the above relations and cannot be inverted—the system defined by \( L_0 \) is, thus, maximally singular. The conjugate pairs \((\omega, \Pi), (\epsilon, \pi), (B, \Phi), (\beta, \phi)\) satisfy the canonical commutation relations (c.c.r.):

\[
\{\omega^A(x), \Pi_C^B(y)\} = \delta^A_C \delta^B_D \delta^D_{xy}, \quad \{\epsilon^A(x), \pi^B(y)\} = \delta^A_B \delta^B_{xy}, \quad \{\beta^A(x), \phi^B(y)\} = \delta^A_B \delta^B_{xy}.
\]

One then constructs the Hamiltonian, schematically

\[
H(q, p) = \frac{1}{2} \dot{q} \dot{p} - L(q, \dot{q}, p) = \phi \dot{q} + H_c(q, p),
\]

following Dirac [47], sometimes also called ‘total’ in order to distinguish it from the ‘canonical’ part \( H_c \), which does not contain primary constraints \( \phi \). It is straightforward to verify that \( H_c \) is indeed explicitly independent of the velocities \( \dot{q} \), which enter as undetermined functions in front of \( \phi \). By this rigorous procedure one gets for the Hamiltonian density:

\[
\mathcal{H} = \Pi'_{AB} \dot{\omega}^A_B + \Pi'_{AB} \dot{\phi}^A_B + \phi'_{AB} \dot{\beta}^A_B - \mathcal{L} = \left( \Pi'_{AB} - \frac{1}{2} \epsilon^{ijk} B_{jAB} \right) \dot{\omega}^A_B + \left( \Pi'_{AB} - \frac{1}{2} \epsilon^{ijk} \beta_{jAB} \right) \dot{\phi}^A_B + \phi'_{AB} \dot{\beta}^A_B + \phi'_{AB} \dot{\beta}^A_B - \mathcal{L}.
\]

The canonical part simply consists of spatial components of Lagrangian \( \mathcal{H} = -\mathcal{L} \mid_{\omega = \pi = 0} \). It contains no momenta whatsoever at this stage, due to primary constraints, and we retain full covariance w.r.t. Lorentz indices. The form of \( \mathcal{H} \) will be specified shortly, after reduction is made (compare with the detailed expressions unfold in the Hamiltonian analysis of closely related BFCG theory [48]).

Next we calculate the development of the primary constraints \( \dot{\phi} = \{\phi, H\} \equiv \chi \) in order to find out the additional consistency requirements for them to preserve in time—the secondary constraints \( \chi \approx 0 \). It may happen that some combinations of constraints form a 2nd class (sub) system, failing to commute. This is precisely our situation, since

\[
\left\{ \Phi_{AB}, \Pi^{CD} - \frac{1}{2} \epsilon^{CDE} B_{DE} \right\} = \frac{1}{2} \epsilon^{ijk} \phi_A \phi_B \phi^C \epsilon^{ijk} B_{jk}, \quad \left\{ \phi_A, \Pi_{AB} - \frac{1}{2} \epsilon^{CDE} B_{DE} \right\} = \frac{1}{2} \epsilon^{ijk} \phi_A \phi^C \epsilon^{ijk} B_{jk}.
\]

Note that the second class nature of the initial Lagrangian \( L_0 \) is not the specialty of our Poincaré modification but is common to any BF theory in various spacetime dimensions. This apparent fact of the full-fledged Dirac’s generalized Hamiltonian analysis is often overlooked in the canonical description of BF and related theories [2, 37, 43, 49] (with rare notable exceptions, e.g. [48, 50]).
One can reduce the system by solving the second class constraints as strong equations. The formal procedure includes passing to the Dirac brackets in order not to sum over variables, which have been thrown away (see [50]). In our case the constraints are of special type, such that we can make a shortcut and simply solve for the spatial $B$ and $\beta$ components (together with their identically vanishing momenta), since these just serve the purpose of identifying (the spatial part of) the $\pi$-connection’s conjugate momenta. The rest of the canonical commutation relations are unaltered, as can be easily verified, and one is left in (4.20) with the last (the spatial part of) the conjugate momenta. The rest of the canonical constraints are first class and all commute among themselves. They give rise to the secondary constraints

$$\{\Pi_{AB}, H\} \equiv \chi_{AB}^0 = D_i \Pi_{AB} - e_i[A] \pi_B^* \approx 0,$$

$$\{\pi_A^i, H\} \equiv \chi_A^0 = D_i \pi_A^i \approx 0,$$

$$\{\Phi_{AB}^0, H\} \equiv \chi_{AB}^0 = e^{i[A]} \omega_{iCB} \approx 0,$$

$$\{\phi_A^0, H\} \equiv \chi_A^0 = e^{i[A]} \omega_{iA} \approx 0.$$  

The derivative $D_i$ is taken w.r.t. the spatial connection $\omega_{iA}$. The canonical part of the Hamiltonian takes form

$$-\mathcal{H}'_c = \omega_0^B \chi_{AB}^0 + e_i^A \omega_{iA}^0 + 2 B_{0i}^A \chi_{AB}^0 + 2 \beta_0^A \phi_i^0 - \partial_i (\omega_0^B \Pi_{AB} + e_i^A \pi_B^*).$$  

The bulk contribution to $\mathcal{H}'$ vanishes as the sum of (the primary as well as secondary) constraints. We did not specify any form of the boundary conditions and kept the surface term explicit. A good cross check is the consistency between the Hamiltonian $\mathcal{H}' = \{ f, H' \}$ and the Lagrangian (4.21) equations of motion, once the solution to second class constraints is taken into account. We warn the reader not to discard the primary constraints from the outset. Although the present case of reduction is very simple, in general, it may affect the symplectic structure of the original action. Moreover, the primary constraints are essential for the equivalence between the Lagrangian and Hamiltonian formulations. By keeping only the (reduced) canonical part $H'$, we cannot even address the gauge transformations on the full phase-space—only the spatial ones.

\[\text{Class. Quantum Grav. 35 (2018) 235007} \]
The completion of Dirac’s procedure consists in proving that $\dot{\chi} \approx 0$ are conserved. This follows from the closure of the algebra:

$$\left\{ \xi_{AB}^0, \chi_{CD}^0 \right\} = \chi_{CD}^0 \eta_{AB}^0 - \chi_{DA}^0 \eta_{BC}^0,$$

$$\left\{ \chi_{AB}^0, \chi_{CD}^0 \right\} = \chi_{AB}^0 \eta_{CD}^0,$$

$$\left\{ \xi_{AB}^0, \chi_{0C}^0 \right\} = \chi_{AB}^0 \eta_{0C}^0,$$

$$\left\{ \chi_{0A}^0, \chi_{0C}^0 \right\} = \chi_{0A}^0 \eta_{0C}^0,$$

$$\left\{ \chi_{0A}^0, \eta_{AB}^0 \right\} = -\chi_{AB}^0.$$

(4.24)

The rest of the commutators being trivially zero.

It is worth at this point to perform the physical degrees of freedom count, in order to make sure that the theory is indeed topological. Starting from the initial phase space of dimensionality (which we denote by putting variables in brackets)

$$[\omega_a^AB] + [\Pi_{AB}] + [\pi_i^A] + [\Phi_{AB}] + [\phi_{ab}] = 200,$$

we eliminate some variables through the strong second class equalities

$$[\Phi_{AB}] + [\phi_{AB}] + [\phi_{0}] = 60.$$

Finally, we perform the symplectic reduction as follows: put the system on the surface of first class constraints, then gauge away the redundant modes by factoring out the action of the gauge generator (for arbitrary functions of time $\epsilon(t)$)

$$G(t) = \sum_{n=0}^{N} \sum_{\alpha} \varepsilon^{(n)}_\alpha \psi^{(N-n)}_\alpha,$$

$$\varepsilon^{(n)}_\alpha = \frac{d^n}{dt^n} \varepsilon_\alpha,$$

(4.25)

through the chains of first class constraints, unambiguously constructed once the set of primary ones (first class) $\{ \alpha \}$ is given. The multi-index $\alpha$ is linked to the tensorial nature of transformations, while $(N - n)$ gives the generation number (primary/secondary/tertiary/etc.). As a by-product, knowing the derivative order of gauge transformations, one can predict the overall number $N$ of generations of constraints, and vice versa.

The chains $G^{(N-n)}_\alpha$ in (4.25) are constructed iteratively as follows:

$$G^{(0)}_\alpha = \text{primary},$$

$$G^{(1)}_\alpha + \left\{ G^{(0)}_\alpha, H \right\} = \text{primary},$$

$$\vdots$$

$$G^{(N)}_\alpha + \left\{ G^{(N-1)}_\alpha, H \right\} = \text{primary},$$

$$\left\{ G^{(N)}_\alpha, H \right\} = \text{primary}.$$

(4.26)
In the present situation the primary ones are
\[ G^0_{(0)AB} = \Pi^0_{AB}, \quad G^0_{(0)L} = \pi^0_L, \quad G^0_{(0)AB} = \Phi^0_{AB}, \quad G^0_{(0)A} = \phi^0_A, \] (4.27)
and the procedure terminates already at the secondary \( n = 0, 1 \):
\[ G^0_{(1)}(x) = -\chi^0(x) + \int d^3y \mathcal{A}^{0\beta}(x, y) \phi^\beta(y). \] (4.28)
where coefficient kernels \( \mathcal{A}^{0\beta} \) are fixed by the last requirement in (4.26) to close onto the primary constraint surface (we have the identical zero due to commutation \( \{ \phi, \phi \} = \{ \phi, \chi \} = 0 \)).

Straightforward calculation gives the total (smeared) generator
\[ \mathcal{G}(\mathcal{U}, u, \Xi, \xi) = -\mathcal{J}(\mathcal{U}) - \mathcal{P}(u) + \mathcal{F}(\Xi) + \mathcal{F}(\xi) \] (4.29)
as a combination of elementary ones:
\[ \mathcal{J}(\mathcal{U}) = \int \dot{u}^A \Pi^0_{AB} - \dot{u}^B (2\omega_{04}^{A} \Pi^0_{CB} - \epsilon_{04}^{A} \pi^0_B + 4B_{04}^{A} \Phi^0_{CB} - 2\beta_{04}^{A} \phi^0_B + \chi^0_{AB}), \]
\[ \mathcal{P}(u) = \int u^A \pi^0_A - u^A (\omega_{00}^{A} \pi^0_B + 2\beta_{00}^{A} \phi^0_B + \chi^0_A). \]
\[ \mathcal{F}(\Xi) = \int \dot{\xi}_i^{AB} \Phi^0_{AB} - \dot{\xi}_i^{0} (2\omega_{00}^{A} \Phi^0_{CB} + \chi^0_{AB}). \]
\[ \mathcal{F}(\xi) = \int \dot{\xi}_i^{AB} \Phi^0_{AB} - \dot{\xi}_i^{0} (\omega_{i0}^{AB} \phi^0_B + \beta_{i0}^{AB} \phi^0_B + \chi^0_{AB}). \] (4.30)

This generalizes the result for the canonical gauge generator of \( SO(3, 1) \) BF theory, reported in [50]. The construction provides the correct transformation properties via
\[ \delta f = \{ f, \mathcal{G} \}. \] (4.31)
mapping solutions into solutions (gauge symmetry). Unlike the secondary constraints (4.22), it acts on the full phase-space of the theory:
The correct covariant expressions (4.8) for all the Lagrangian field components (spatial as well as temporal; using also the 2nd class relations \( \delta B_0^B = \epsilon_{ijk} \delta B_1^C \delta B_2^D \)) are reproduced within the Hamiltonian framework, thus exhibiting the equivalence between the two pictures.

Using the Jacobi identity, the commutator between the two consecutive transformations is given:

\[
(\delta_1 \delta_2 - \delta_2 \delta_1) f = \{ \{ \mathcal{G}_1, \mathcal{G}_2 \}, f \}. \tag{4.32}
\]

Its elementary constituents realize the generalized matrix commutators:

\[
\begin{align*}
\{ \mathcal{J}(\mathcal{U}_1), \mathcal{J}(\mathcal{U}_2) \} &= -\mathcal{J}(\mathcal{U}_1 \mathcal{U}_2), \\
\{ \mathcal{J}(\mathcal{U}), \mathcal{P}(u) \} &= -\mathcal{P}(\mathcal{U} \triangleright u), \\
\{ \mathcal{J}(\mathcal{U}), \mathcal{F}(\Xi) \} &= -\mathcal{F}(\mathcal{U} \triangleright \Xi), \\
\{ \mathcal{P}(u), \mathcal{F}(\xi) \} &= +\mathcal{F}(u \triangleright \xi), \\
\{ \mathcal{P}(u), \mathcal{F}(\Xi) \} &= -\mathcal{F}(u \triangleright \Xi), \\
\{ \mathcal{P}(u), \mathcal{P}(v) \} &= -\mathcal{P}(u \triangleright v) - \mathcal{P}(v \triangleright u).
\end{align*}
\]

We use a chance to comment here on the relation between the dynamics and gauge in reparametrization invariant systems (see ‘the problem of time’). The Hamiltonian—generator of time evolution—in such model is a combination of first class constraints, which are also known to generate the gauge transformations (i.e. ‘unphysical’ changes in the description of the system). Working on the full phase space allows to disentangle these notions: the specific combinations of first class constraints are different for two objects \( \mathcal{H} \) and \( \mathcal{G} \); the key role is played by the primary set.

One usually defines the notion of Dirac observables w.r.t. individual constraints \( \{ \chi, f \} = 0 \) (often disregarding the primary \( \phi \approx 0 \), working on the smaller phase space). In the quantum theory, one represents the canonical variables via operators on the appropriate Hilbert space \( \mathcal{H} \) of states of the system, the Poisson (Dirac) brackets being replaced by a commutator \([. , . \]) = i{ , . \}). For instance, in our example:

\[
\begin{align*}
\mathcal{J}(\mathcal{U}) &\mapsto \mathcal{U} \cdot \mathcal{J}, \\
\mathcal{P}(u) &\mapsto u \cdot \mathcal{P}
\end{align*}
\tag{4.33}
\]

are the elements of the (local) Poincaré algebra (4.10). The Dirac prescription then consists in imposing on states \( \chi_\alpha |\Psi\rangle = 0 \) individually for each \( \alpha \), which are then consistent for the first class system.

With the distinction just pointed out between \( \mathcal{H} \) and \( \mathcal{G} \) on the full phase space, the function of canonical variables may satisfy two \textit{a priori} distinct conditions: \( \{ \mathcal{H}, F \} = 0 \) and/or \( \{ \mathcal{G}, F \} = 0 \). The first can be thought of as characterizing ‘evolving constants of motion’, i.e. uniquely associated with the state—solution of the e.o.m. The state itself, however, is not uniquely defined by the Cauchy data and depends on the arbitrary functions, entering the Hamiltonian (in our reduced case, the velocities of \((\epsilon, \omega, \beta, B)\) time components, associated with the 1st class primary constraints on momenta, are not defined by the evolution equations). Thus, the first type functionals may depend on the gauge choice for particular Hamiltonian, whereas the second condition then characterizes ‘gauge-invariant’ functionals.

Requiring the time preservation of vanishing of currents \( \mathcal{G}^\alpha \) on every hypersurface, by construction we have then \( 0 = \partial_t \mathcal{G}^\alpha = \{ \mathcal{H}^\alpha, \mathcal{G}^\alpha \} \), and the state has to satisfy Hamiltonian e.o.m. So that symmetry generators provide an example of the first type functionals. The commutation relations (4.32) express the fact that the canonical (pre-)symplectic structure is degenerate on the constraint surface. One then passes to the quotient w.r.t. the gauge direction, considering the gauge equivalent classes of solutions as ‘physical’ states. The construction of
the appropriate observable algebra is of primary importance for the quantization, especially in gravitational theories, so the Dirac’s ‘rule of thumb’ for all 1st class constraints should be applied with certain care.

5. Poincaré–Plebański (re)formulation of GR

Now, when we have the frames \( e \) at our disposal among the legitimate dynamical variables, it is straightforward to implement the simplicity of the bivectors, in order to reproduce gravity sub-sector. Multiple choices of how to do this are conceivable. First of all, one can simply replace \( B \rightarrow \ast e \wedge e \) directly in the action integral:

\[
S[e, \omega, \beta] = \int_\mathcal{M} \frac{1}{2} \epsilon_{ABCD} e^A \wedge e^B \wedge F^{CD} + \beta_A \wedge T^A. \tag{5.1}
\]

Secondly, one could try to achieve the same effect via the Lagrange multipliers approach, imposing the simplicity in its most direct sense:

\[
S[B, \varpi, \Lambda] = S_0[B, \varpi] + \int \Lambda^{AB} \wedge (\epsilon_{ABCD} e^C \wedge e^D), \tag{5.2}
\]

with the free independent multiplier two-forms \( \Lambda \).

It turns out that the simplicity constraints can also be put into form, linear in both \( B \) and \( e \), which is more in the vein of current spin foam models. In order to stay self-contained and explicit, let us formulate the following

**Lemma (Linear simplicity).** Provided that the tetrad field is non-degenerate, and hence the map \( e \) is invertible, the bivector field \( B \) is simple if and only if either of the two equivalent sets of constraints is satisfied:

\[
\begin{cases}
    *B_{ab}^b c_e = 0 \forall c \notin \{a, b\}, \\
    *B_{ab}^b c_e = *B_{ab}^d c_b c_d = 0 \forall a \notin \{b, c, d\}, \\
\end{cases}
\]

\[
\Leftrightarrow \begin{cases}
    B_{AB} e^B e_c = 0 \forall c \in \{a, b\}, \\
    B_{AB} e^B e_a (b e^c) = 0 \forall c \notin \{a, b\}. \\
\end{cases} \tag{5.3}
\]

**Proof.** First, it is straightforward to verify that the two systems imply each other by noting that for some fixed \( a \neq b \):

\[
*B_{ab}^b c_e = \sum_{cd} \frac{1}{2} \epsilon_{abcd} B_{cd} e^B e^c = \epsilon^{abcd} B_{AB} e^B e^a \quad \text{(no sum over [cd] \neq [ab]).}
\]

Assuming the conditions on the one side, the other one follows from here. Notice also that in our notation the dual becomes \( *B_{ab}^b = (\Sigma^{-1})_{ab} \) on the constrained surface, where \( \Sigma = e \wedge e \).

To show the necessity of these conditions for simplicity of bivectors, it is enough to cast \( B = *e \wedge e \) into form\(^{13}\)

\[
B_{AB} e^B e_c = \epsilon_{ABCD} e^B e_a e^d e^c = (\det e) e^A \epsilon_{dabc}. \tag{5.4a}
\]

or, equivalently:

\(^{13}\) Having in mind the possible interpretation in terms of discrete geometry, (5.4a) can be suggestively referred to as ‘pyramid’, or ‘three-volume’ form of simplicity constraints—namely, by the nature of the object appearing on the right, and since the whole formula can be read as the expression for the volume of a pyramid with the base \( \Sigma \) and a height \( e \).
\[ *B_{c}^{\alpha \beta} = (\text{det} \, e) (e_{c}^{\alpha} \delta^{\beta}_{\epsilon} - e_{c}^{\beta} \delta^{\alpha}_{\epsilon}). \]  
\hspace{1cm} (5.4b)

Observing that the l.h.s. is already linear leads us to the (dual) analogue of ‘cross-simplicity’ constraint in the first line of (5.3), when the r.h.s. in (5.4) is zero. In turn, the non trivial expression on the right in (5.4) restricts the l.h.s. in (5.4a) to be totally antisymmetric in \([abc]\), whereas it is independent of \(b = c \neq a\) in (5.4b), leading to the second line of (5.3), respectively (no sum over spacetime indexes).

In order to demonstrate that the conditions (5.3) are also sufficient, one follows the same reasoning as in [13]. Namely, the generic bivector field can be expanded over the basis, spanned by the skew-symmetric products of \(e\):

\[ *B_{\alpha \beta} = G_{\alpha \beta}^{[cd]} e_{\epsilon}^{[c]e_{d}], \]  
\[ G_{\alpha \beta}^{[cd]} = G_{\alpha \beta}^{[ac]} [A^e B^d] , \]

which after substitution into the first line of (5.3) leads to

\[ *B_{\alpha \beta} = G_{\alpha \beta}^{[cd]} e_{\epsilon}^{[c]e_{d}] } (\text{no sum over } \alpha \beta) \]

The individual normalization coefficients have to satisfy symmetry \(G_{\alpha \beta} = G_{\beta \alpha}\), \(G_{\alpha \alpha} = 0\), but apart from that can be arbitrary. It is only after substitution of this ansatz into the second line of (5.3) that we get the restriction

\[ G_{\alpha \beta} = G_{\alpha \delta} = G_{\delta \alpha} \ \forall \alpha \notin \{b, c, d\}, \]

leading to the equality among all \(G\)'s. Thus, the \(B\) is simple up to an overall factor, which can be eaten by appropriate normalization.

The continuous formulation that we are advocating for is somewhat different from that of Gielen–Oriti’s linear proposal [13], which uses three-forms \(\vartheta = *e \wedge e \wedge e\), and bivectors \(\Sigma = e \wedge e\) as independent variables, but rather represents its dual version. Before introducing the action principle, and in order to make closer contact between the two formulations, we first recall the corresponding constraint term in [13] and notice that this can be rewritten as

\[ \int d^4x \, \Xi_{\alpha \beta}^{[ab]} \, \Sigma_{\beta cd} \, \delta_{A} \vartheta_{Bcde} = \int d^4x \, \Xi_{\alpha \beta}^{[ab]} \, \Sigma_{\beta cd} \, \tilde{\vartheta}_{B}. \]  
\hspace{1cm} (5.5)

One can choose to work either with three-forms (4.1) or, equivalently, their dual densitized vectors:

\[ \tilde{\vartheta}_{A}^{\alpha} = \frac{1}{3!} \epsilon^{abcd} \vartheta_{A}^{abcd} = (\text{det} \, e_{B}^{\alpha}) e_{A}^{\alpha}. \]  
\hspace{1cm} (5.6)

Correspondingly, the Lagrange multipliers \(\Xi_{\alpha \beta}^{[ab]}\) should be exact tensors, s.t.

\[ \Xi_{\alpha \beta}^{[ab]} = \frac{1}{3!} \epsilon^{abcd} \Xi_{\alpha \beta}^{[cd]} - \delta_{\alpha}^{[a} \delta_{\beta}^{b]} \text{– tensor densities.} \]

The somewhat convoluted index symmetries that \(\Xi\) has to satisfy can be restated as the traceless condition on \(\Xi_{\alpha \beta}^{[ab]} = 0\), which upon variation then leads to the appearance of non-trivial Kronecker deltas on the right:

\[ \delta \Xi \Rightarrow \tilde{\vartheta}_{A}^{\alpha} \Sigma_{\alpha \beta}^{[ab]} = \delta_{A}^{[a} \delta_{\beta}^{b]} \text{ for some } \vartheta_{b}. \]  
\hspace{1cm} (5.7)

The antisymmetry in \([AB]\) and the tensorial nature of \(\Sigma\) leave us no choice other than \(\vartheta_{A}^{\alpha} \propto \tilde{\vartheta}_{A}^{\alpha} = e_{\alpha}^{A}\), and we get the simplicity up to an overall normalization, which is irrelevant. Applying the Hodge dual \(*\), one restates this in terms of three-forms, resulting from variation w.r.t. \(\Xi\), correspondingly:

\[ \Sigma_{\alpha \beta}^{[ab]} \vartheta_{Bcde} = \vartheta_{A}^{\alpha} \epsilon_{Bcde} - \vartheta_{B}^{\alpha} \epsilon_{Acdde}, \]  
\hspace{1cm} (5.8)
which are essentially the original Gielen–Oriti’s constraints.

Comparing (5.7) with (5.4b), and juxtaposing them against the constraint term (5.5), then suggests the respective least action principle in terms of dual variables $B \leftrightarrow \Sigma$ and $e \leftrightarrow \vartheta$, correspondingly:

$$S_{pp}[\mathcal{B}, \varpi, \Theta] := S_0[\mathcal{B}, \varpi] + \int \left( \Theta^A \delta e^B \right) \wedge B_{AB}, \quad (5.9)$$

which we coined, referring to its gauge group, the Poincaré–Plebański formulation (although such a name might be as well attributed either to the ‘A-version’, or essentially to any formulation of this flavour). The $4 \times 4 \times 6 = 96$ Lagrange multipliers constitute the tangent $T\mathcal{M}$-valued two-forms, that is

$$\Theta^A = \frac{1}{2} \Theta_{ab}^A \, dx^a \wedge dx^b \otimes \partial_c, \quad (5.10)$$

are the sections of the fiber bundle $\bigwedge^2 T^*\mathcal{M} \otimes T\mathcal{M}$. In the constraint term of the action (5.9) they contract with tetrad one-forms using the pairing $dx^a \otimes \partial_b = \delta^a_b$ in the tangent vector index: $\Theta^A \delta e^B = \Theta_{ab}^A \, dx^a \wedge dx^b \otimes \partial_c$. The $\Theta^A$’s are restricted to be traceless $\Theta_{ab}^{ab} = 0$, that is possess the components of the form $\Theta_{ac} + \frac{1}{2} \delta_c^a \Theta_{bd}^{bd}$, we can formulate this in the coordinate independent way as the full contraction with the canonical tangent-valued form on $\mathcal{M}$ being zero:

$$\theta_{\mathcal{M}} := dx^a \otimes \partial_a, \quad \theta_{\mathcal{M}}, \Theta^A = 0. \quad (5.11a)$$

Let’s count the number of independent $\Theta$ components, in order to verify that we have enough of them to eliminate $36 \, B_{cd}^{BD}$ in favour of $16 \, e_{ac}^A$. Apart from the $4 \times 4$ traceless conditions (5.11), from the contraction with $B$ in the action (5.9) follow $10 \times 6$ antisymmetrization equations

$$\Theta^{(A \delta e^B)} = 0, \quad (5.11b)$$

which $\Theta$ and $e$ have to satisfy. Subtracting from this 16 d.o.f. corresponding to $e$’s (they just serve the purpose to isomorphically map indices $e(x) : T_x\mathcal{M} \rightarrow \mathbb{R}^{3,1}$), we are left with $60 - 16 = 44$ additional requirements on $\Theta$. Thereby we get the total number of $96 - 16 - 44 = 36$ independent $\Theta$’s—exactly the right amount to enforce simplicity.

It should be more or less evident after our exposition that the symmetries of the Lagrange multipliers lead to the variation, constrained by the system (5.3), depending whether we choose to vary w.r.t. $\Theta$ or its dualized version $\bar{\Theta}$. The lemma then implies that this is the same as performing variation on the simplicity constraint surface. The manifest presence of the Hodge-star $\ast$ in constraint (5.2) becomes shrouded, instead one has the restriction on the multipliers $\Theta$. The free variation of $\delta \Lambda$ equates the constraint pre-factor to zero exactly, whilst for $\delta \Theta$ obeying additional conditions—we get the non-vanishing expression on the r.h.s. In an analogous situation within the standard Plebański quadratic approach, the corresponding quantity on the right is usually interpreted in geometric terms as a definition of the four-volume (on the solution of constraints), whilst a non-trivial symmetrization conditions are put on the l.h.s. It is these latter conditions that actually constitute the substance of the respective ‘volume’ part of simplicity constraints. They require that the definition of the four-volume be consistent, i.e. does not depend on the multiple choices that could be made for its parametrization on the l.h.s. Note that in (5.4) we get the very same picture, now with the quantity on the r.h.s. being precisely the non-trivial three-volume (see (3.27)). At the same time this last bit now is ‘localized’ at the level of each tetrahedron, irregardless of the whole four-simplex, which was the case for quadratic version (iii). Lastly, the analogue of the ‘cross-simplicity’,
when the r.h.s. is zero, now expresses that the corresponding (discrete) \( e \) is collinear with the face \( S_f \), being orthogonal to its dual bivector \( B_f = \ast \Sigma_f \).

These three \textit{a priori} distinct choices for constraint imposition, tabulated above, all seem to represent the same physical content. In either of the \( \Lambda \) or \( \Theta \) versions, inserting further the solution for \( B \) back into action, one reduces the initial topological theory (4.4) to that of (5.1), that is the Einstein–Cartan action (2.2) supplemented with an extra term for (zero) torsion. At first sight, this might seem an excess, since the variations \( \delta \omega \) of the EC-term alone incidentally give the vanishing of \( De = 0 \) on-shell. However, the two theories are not identical: \( \beta \) plays the role of the Lagrange multiplier imposing the torsionless constraint \( T = 0 \), which by the rule of procedure is required when we pass from the 2nd to 1st order formulation, in order to preserve the original dynamical content of the theory. Thus, we expect the equivalence should hold with the Einstein–Hilbert variational principle, and not with the ‘Palatini variation’ method.

The relations between different action principles can be schematically depicted in a diagram:

\[
\begin{align*}
1\text{st order:} & \quad S_{\text{PP}}[\mathcal{B}, \varpi, \Theta] \xrightarrow{\delta \Theta} S_{\text{GC}}[e, \omega] + \int \beta \wedge T \\
2\text{nd order:} & \quad S_{\text{GO}}[B, e, \omega[e], \Theta] \xrightarrow{\delta \Theta} S_{\text{GC}}[e, \omega[e]] \equiv S_{\text{EH}}[e]
\end{align*}
\]

In the bottom left corner appears a variant of the ‘hybrid’ action of the form dual to that of Gielen–Oriti [13], but with the unique \( e \)-compatible torsion-free spin connection. This is to be contrasted with their 1st order formulation, where \( \omega \) is independent and the gauge status of non-dynamical \( \vartheta \sim e \) is less clear, which enters a separate sequence:

\[
S_{\text{PP}}[\mathcal{B}, \varpi, \Theta] + \int \mu \wedge \beta \xrightarrow{\delta \mu} S_{\text{GO}}[B, e, \omega, \Theta] \xrightarrow{\delta \Theta} S_{\text{EC}}[e, \omega].
\]

We stress that the reduction of the Einstein–Cartan theory to that of GR is achieved only \textit{on-shell} in vacuum, by solving the dynamical e.o.m. for \( \omega \). In contrast, one puts additional restrictions on the allowed variations of the generalized coordinates by the use of (non-dynamical) Lagrange multipliers \( \Theta, \beta \), which then acquire the physical meaning of ‘reaction forces’, corresponding to variations that violate the constraints. The discussions of the relation between two approaches have been recurrent in the literature in the past, in particular, regarding the higher order Lagrangians and matter couplings (e.g. see [52] and references therein). It is a firmly established fact that the constrained variations should lead to the same result as for the case where constraints have been already solved from the outset. On the other hand, by allowing arbitrary variations w.r.t. d.o.f. which were previously restrained to lie on the constrained surface\(^{14}\) the presumed ‘equivalence’ with the original (2nd order) set-up is skewed. One plainly does not possess the same d.o.f. in two approaches to variation. This might be of relevance for the precise form of dynamical symmetries, by the Noether’s theorem, since the variations would in general contain terms off the constraint surface.

6. Conclusions and outlook

In the first part of this work we reviewed the classical Plebański formulation of gravity, which underlies current spin foam models of EPRL and FK. We considered both the quadratic version and the fully linear formulation with three-forms (4D normals), paying a special attention

\(^{14}\) Often referred to as ‘Palatini variation’—though somewhat erroneously (see discussion in [53]).
to the implementation of the ‘volume’ part of simplicity constraints. Our revision of its quadratic version in the symmetry reduced setting of cuboids revealed that one cannot replace it with the 3D closure condition, in general, contrary to triangulations. As result, there is no unique geometric four-volume. The linear case puts non-trivial conditions on normals and bivectors, which ensure the existence of edges/matching of shapes/uniquely defined three-volumes, in quite the intricate way. This prompts us to pass to the tetrad variables, instead of normals.

In the second part of this work we considered the modification of the classical action principle by putting torsion to zero. To explore the consequences we studied in section 4 the corresponding change in the BF theory and its larger gauge group—the Poincaré (affine) extension of the (homogeneous) Lorentz group—developing our analysis in detail both at the Lagrangian and Hamiltonian levels. Our Dirac’s generalized constraint analysis, in fact, shorts the derivation in [48], corroborating also the equivalence with covariant framework by an actual construction of canonical gauge generator. The extended action presents a perfect ground for the imposition of simplicity constraints, as we have frames explicitly at our disposal. In section 5 we present an alternative look at the linear simplicity constraints, dual to that of [13], and comment on the relations between various formulations.

One point to be noticed is that in our reformulation it becomes apparent that the tetrad field by and large presents in the EPRL model right away. It sneaks in disguise of a four-vector normal to the tetrahedra, required for linear reformulation of constraints. Recalling the original Plèbanski’s quadratic constraints, this was based on the appearance of tetrad in the action only in a specific combination, which could be collectively denoted \( B = \star e \wedge e \), and has to satisfy some algebraic relations. The same situation is encountered in LQG with the conjugate momenta and reality conditions [54]. However, let us stress that the role of soldering form as an independent entity is much more than that: it encodes all the metric properties and is directly related to the diffeomorphisms, playing the part of the gauge potential of local translations. These roles of tetrad are hardly appreciated in a formulation, where it is basically excluded from consideration as a configuration variable, and diminished to just the momenta. In essence, upon a closer look the theory is indistinguishable from the classical Einstein–Cartan, or Einstein–Hilbert gravity (leaving aside the Barbero–Immirzi parameter), and one has to face the task of directly handling tetrads at the discrete and quantum levels\(^{15}\).

Despite the classical nature of our results, it is clear from the context in which we put the present work, that we expect the view developed here to be of relevance for quantum gravity, both at the level of path integral (SF) and for covariant canonical loop-quantization. In particular, the Plebański approach applied to the BF-Poincaré theory displays minimal distinction from the correspondingly constrained 1st order formulation, thus reinforcing one’s expectations for a better contact between the two quantization programs. As an outlook, let us observe several issues that one may face on this route.

- The choice of an appropriate discretization for a frame field and constraints. The guidance may be provided by a twofold nature of the frame field:

  - On the one hand side, the vector \( V = \mathbb{R}^{3,1} \)-valued one-form \( e \), being a gauge potential of translations, combines naturally with the homogeneous Lorentz part \( \omega \in \mathfrak{h} \) into a single Cartan connection one-form \( \varpi = \omega + e \), taking values in larger algebra \( g \). Treating both parts on the same footing—at least in the topological BF case—thus suggests a corresponding discretization in terms of generalized \( G \)-holonomies: besides the parallel transport along paths in \( \mathcal{M} \) (given by the Ehresmann \( H \)-connection), a Cartan con-

\(^{15}\) The similar views were recently advocated [55], in the context of canonical 3D LQG.
connection gives also a notion of ‘development on the model $G/H$-Klein geometry’ (the ‘rolling’, or translation of the point of tangency; see [45] and references therein)\(^{16}\).

– On the other side, in the constrained case of gravity $e$ provides the basis for the geometric ‘simple’ $n$-forms (and for the $\omega$’s conjugate momenta two-forms, in this way). In the end, one could expect similarity of the discretization with the variant of Regge calculus that comes from the gauge theoretic approach to gravity [56, 57]. The discrete $e$-field may appear in the ‘integrated’ form (i.e. conjugated by the $H$-holonomies) in order to ensure gauge invariance. The dual form (5.4a) of simplicity constraints associates naturally a 3-volume normal vector to every boundary polyhedron, so the generalization to a higher valence case seems to be within the reach (e.g. using some variant of Minkowski’s theorem).

\* It is sensible to first gain some experience with the SF/loop quantization of the respective topological BF theory. The $H$-holonomy encodes the information about connection up to transformations, leaving source and target intact. Note that one gets an element of $G \supset H$ for the lift of a curve w.r.t. Cartan connection $g \cong h \oplus V$, where the latter isomorphism should hold as $ad_H$-representations, for reductive Cartan geometries. The space of connections is now significantly larger, so that $H$-transformations cannot reduce the gauge d.o.f. sufficiently.

Allowing for more gauge transformations could deal with this issue, and indeed—the full $ad_G$ symmetry (4.12) suggests to use the basis of $G$-invariant spin networks in the topological BF case. Note that $\varpi$ is essentially Ehresmann connection in the $G$-bundle, corresponding to vertical automorphisms. However, the reduction to gravity attaches the affine frames to the manifold, while soldering the $\mathbb{R}^{3,1}$ part to the tangent directions, corresponding to (horizontal) diffeomorphisms. The connection here is Cartan’s absolute parallelism (on $H$-bundle). This distinction seems to be crucial, and the usual procedure should be specialized correspondingly.

**Acknowledgments**

The author is grateful to Benjamin Bahr and Sebastian Steinhaus for stimulating discussions and for useful comments on an earlier draft of this paper. This work was funded by the project BA 4966/1-1 of the German Research Foundation (DFG).

**Appendix A. Conventions and notation**

We use the Latin letters from the beginning of the alphabet to denote the covariant field components: lowercase $a, b, c, \ldots = 0, 1, 2, 3$ for the world tensors w.r.t. the holonomic coordinate basis, and capital $A, B, C, \ldots = 0, 1, 2, 3$ for the (internal) Lorentz coordinates w.r.t. the orthonormal locally inertial frames. For the $3 + 1$ space/time split, the letters from the middle of the alphabet are used $i, j, k, \ldots = 1, 2, 3$ for the spatial field components.

Anti-symmetrization of indices is performed with the respective order $|S| = n!$ of the symmetry group in the denominator, s.t the projection property holds, and denoted by the square brackets:

\(^{16}\) We assume, an incorporation of this notion might pave the way for a better control over diffeos in the discrete and quantum gravity.
\[ t_{[a_1 \ldots a_n]} := \frac{1}{n!} \sum_{\pi \in S_n} \text{sign}(\pi) t_{a_{\pi(1)} \ldots a_{\pi(n)}}. \quad (A.1) \]

The internal \( \mathbb{R}^{3,1} \) comes with the Minkowski metric \( \eta_{AB} = \text{diag}(-1, 1, 1, 1) \) and the totally anti-symmetric Levi-Civita symbol \( \epsilon^{ABCD} = 4! \delta^A_0 \delta^B_1 \delta^C_2 \delta^D_3 \). (Note that, although we prefer ‘mostly plus’ convention for the metric signature, we maintain the full Lorentz covariance and actually never use this explicitly.) The metric \( \eta \) allows to freely raise and lower internal indices, identifying \( \mathbb{R}^{3,1} \) with its dual, e.g. \( \epsilon_{ABCD} = \eta^{AA'} \eta^{BB'} \eta^{CC'} \eta^{DD'} \epsilon^{A'B'C'D'} = -4! \delta^1_0 \delta^2_1 \delta^3_2 \delta^4_3 \).

It defines the (internal) Hodge-star duality operator (in arbitrary dimension \( N \)):

\[
\mathcal{A}_1 \ldots \mathcal{A}_{N-n} \rightarrow \mathcal{A}_1 \ldots \mathcal{A}_N \Rightarrow \star \mathcal{A}_1 \ldots \mathcal{A}_n := \frac{1}{n!} \epsilon_{\mathcal{A}_1 \ldots \mathcal{A}_n} \mathcal{B}_{n+1} \ldots \mathcal{B}_N \mathcal{Q}^{\mathcal{B}_n+1 \ldots \mathcal{B}_N}. \quad (A.2) \]

Analogously, the spacetime totally-skew symbol \( \epsilon^{abcd} \) is defined to have \( \epsilon^{0123} = \epsilon_{0123} = 1 \) for both upper and lower indices. To it corresponds the respective Hodge-star \( \star \) on spacetime exterior algebra. The contraction properties are

\[
\epsilon^{a_1 \ldots a_n b_1 \ldots b_{n+1} \ldots b_0} = n!(N-n)! \delta_{b_1}^{[a_1} \cdots \delta_{b_0}^{a_n]}, \quad (A.3) \]

and similarly for the internal \( \epsilon \), just with the minus sign, in particular \( \epsilon^{ABCD} \epsilon^{ABCD} = -4! \).

Working with \( \epsilon \) allows to conveniently express the determinants, e.g. the elementary 4D volume spanned by the coordinate basis (co)vectors

\[
dx^a \wedge dx^b \wedge dx^c \wedge dx^d = d^4x \epsilon^{abcd}. \quad (A.4)\]

Intuitively, \( \epsilon \) (in conjunction with the frame \( e \)) generalizes the flat Euclidean vector cross-product to arbitrary spacetimes and their subspaces. Thus, we get the dual vectors \( (4.1) \), representing locally the 3D volume normal to the elementary parallelepiped of the hypersurface. Similarly, the simple bivectors \( (2.4) \) from the II-sector represent locally the 2D area normals to the surface’s elementary parallelograms.

We use the \( \mathcal{J} \)-symbol for contractions of tensors’ components, employing the duality of elementary coordinate co/vectors \( dx^a, \partial_b = \delta_b^a \). For example, the internal product of a vector with coordinate basis \( n \)-forms:

\[
\partial_b \cdots (dx^a_1 \wedge \ldots \wedge dx^a_n) = n! \delta^a_b dx^a_1 \wedge \ldots \wedge dx^a_n. \quad (A.5)\]

It is ordinarily clear from the context which of the components of tensors are being contracted.

**Appendix B. On the teleparallel ‘gauge’**

As a side remark, let us touch upon how one can alternatively arrive, starting from the same unconstrained action \( (4.4) \), to the so called ‘teleparallel equivalent of GR’. In place of simplicity for \( B \), one may choose to constrain \( \beta \) in the original Poincaré BF action \( (4.4) \). One can split up the generic metric-preserving connection into \( e \)-compatible (torsionless, Levi-Civita) part and contortion tensor \( K \), respectively:

\[
\omega^{AB} := \omega^{[A} e^{B]} + K^{AB}, \quad T^A := K^A_B \wedge e^B. \quad (B.1)\]

If we require \( \beta \) to be of the form:
\[ \beta_A = \frac{1}{2} \epsilon_{ABCD} \epsilon^B \wedge K^{CD}, \] (B.2)

we obtain the theory of distant parallelism with non-trivial torsion, written in components as follows:

\[ S_\parallel = \int d^4x e \left( \lambda_{AB}^{ab} F_{ab}^{AB} - \frac{1}{4} T_{ab}^{c} T_{ab}^{c} e + \frac{1}{2} T_{ab}^{c} T_{ab}^{c} + T_{ab}^{a} T_{ab}^{b} \right), \quad e \lambda_{AB}^{ab} := \frac{1}{2} e B_{AB}^{ab}. \] (B.3)

The relation to general relativity is established via following identity:

\[ e \left( e_{AB}^{ab} F_{ab}^{AB} - \frac{1}{4} T_{ab}^{c} T_{ab}^{c} e + \frac{1}{2} T_{ab}^{c} T_{ab}^{c} + T_{ab}^{a} T_{ab}^{b} \right) = eR - \partial_a \left( e e_{AB}^{ab} K_{AB}^{ab} \right), \] (B.4)

where \( R \) is the Ricci scalar curvature (built from Levi-Civita connection). That is for the vanishing curvature \( F \) (of the so called Weitzenböck connection), the part quadratic in torsion differs from the Einstein’s theory by a total divergence. The identity (B.4) also demonstrates explicitly the difference between EC and EH Lagrangians.

**ORCID iDs**

Vadim Belov  \( \text{https://orcid.org/0000-0002-3377-4411} \)

**References**

[1] Holst S 1996 Barbero’s Hamiltonian derived from a generalized Hilbert–Palatini action *Phys. Rev. D* **53** 5966–9
[2] Perez A 2013 The spin foam approach to quantum gravity *Living Rev. Relativ.* **16**
[3] Barrett J W 1995 Quantum gravity as topological quantum field theory *J. Math. Phys.* **36** 6161–79
[4] Oeckl R 2003 A ‘general boundary’ formulation for quantum mechanics and quantum gravity *Phys. Lett. B* **575** 318–24
[5] Engle J, Pereira R and Rovelli C 2008 Flipped spin foam vertex and loop gravity *Nucl. Phys. B* **798** 251–90
[6] Engle J, Livine E, Pereira R and Rovelli C 2008 LQG vertex with finite Immirzi parameter *Nucl. Phys. B* **799** 136–49
[7] Freidel L and Krasnov K 2008 A new spin foam model for 4D gravity *Class. Quantum Grav.* **25** 125018
[8] Livine E R and Speziale S 2008 Consistently solving the simplicity constraints for spin foam quantum gravity *Europhys. Lett.* **81** 50004
[9] Alexandrov S and Roche P 2011 Critical overview of loops and foams *Phys. Rep.* **506** 41–86
[10] Barrett A and Oriti D 2012 Group field theory and simplicial gravity path integrals: a model for Holst–Plebanski gravity *Phys. Rev. D* **85** 044003
[11] Livine E R and Dupuis M 2010 Lifting \( SU(2) \) spin networks to projected spin networks *Phys. Rev. D* **82** 064044
[12] Reisenberger M P and Rovelli C 1997 ‘Sum over surfaces’ form of loop quantum gravity *Phys. Rev. D* **56** 3490–508
[13] Gielen S and Oriti D 2010 Classical general relativity as BF-Plebanski theory with linear constraints *Class. Quantum Grav.* **27** 185017
[14] Barrett J W, Dowdall R J, Fairbairn W J, Gomes H and Hellmann F 2009 Asymptotic analysis of the EPRL four-simplex amplitude *J. Math. Phys.* **50** 112504
[15] Barrett J W, Dowdall R J, Fairbairn W J, Hellmann F and Pereira R 2010 Lorentzian spin foam amplitudes: graphical calculus and asymptotics *Class. Quantum Grav.* **27** 165009
[16] Kamiński W, Kiezlowski M and Lewandowski J 2010 Spin-foams for all loop quantum gravity Class. Quantum Grav. 27 095006
[17] Kamiński W, Kiezlowski M and Lewandowski J 2010 The EPRL intertwiners and corrected partition function Class. Quantum Grav. 27 165020
[18] Dona P, Fanizza M, Sarno G and Speziale S 2018 SU(2) graph invariants, Regge actions and polytopes Class. Quantum Grav. 35 045011
[19] Bahr B and Steinhaus S 2016 Investigation of the spinfoam path integral with quantum cuboid intertwiners Phys. Rev. D 93 104029
[20] Freidel L and Speziale S 2010 Twisted geometries: a geometric parametrisation of SU(2) phase space Phys. Rev. D 82 084040
[21] Freidel L, Geiller M and Ziprick J 2013 Continuous formulation of the loop quantum gravity phase space Class. Quantum Grav. 30 085013
[22] Freidel L and Ziprick J 2014 Spinning geometry = twisted geometry Class. Quantum Grav. 31 045007
[23] Dittrich B and Ryan J P 2011 Phase space descriptions for simplicial 4D geometries Class. Quantum Grav. 28 065006
[24] Magliaro E and Perini C 2013 Regge gravity from spinfoams Int. J. Mod. Phys. D 22 1350001
[25] Bi Y 1993 Topological Field Theory and Observables PhD thesis University of New Brunswick
[26] Magliaro E and Perini C 2013 Regge gravity from spinfoams Class. Quantum Grav. 30 165020
[27] Don J W and Crane L 1998 Relativistic spin networks and quantum gravity J. Math. Phys. 39 3296–302
[28] Plebanski J F 1977 On the separation of Einsteinian substructures J. Math. Phys. 18 2511–20
[29] Plebanski J F 1978 On the separation of Einsteinian substructures J. Math. Phys. 19 776–88
[30] Don J W and Crane L 1998 Relativistic spin networks and quantum gravity J. Math. Phys. 39 3296–302
[31] Bianchi E, Dona P and Speziale S 2011 Polyhedra in loop quantum gravity Phys. Rev. D 83 044035
[32] Buffenoir E, Henneaux M, Noui K and Roche P 2004 Hamiltonian analysis of Plebanski theory Class. Quantum Grav. 21 5203–20
[33] Engle J 2011 The Poincaré 2-group and quantum gravity Class. Quantum Grav. 28 045007
[34] Barrett J W and Crane L 1998 Relativistic spin networks and quantum gravity J. Math. Phys. 39 3296–302
[35] Alessandro S 2010 New vertices and canonical quantization Phys. Rev. D 82 044033
[36] Alessandro S 2008 Spin foam model from canonical quantization Phys. Rev. D 77 024009
[37] Alessandro S 2008 Simplicity and closure constraints in spin foam models of gravity Phys. Rev. D 78 044033
[38] Bahr B and Belov V 2017 Volume simplicity constraint in the Engle–Livine–Pereira–Rovelli spin foam model Phys. Rev. D 97 086009
[39] Cuesta V, Montesinos M, Velázquez M and Vergara J D 2008 Topological field theories in n-dimensional spacetimes and Cartan’s equations Phys. Rev. D 78 064040
[40] Wise D K 2010 MacDowell–Mansouri gravity and Cartan geometry Class. Quantum Grav. 27 155010
[41] Kirilchcheva N and Kuzmin S V 2011 The Hamiltonian formulation of general relativity: myths and reality Cent. Eur. J. Phys. 9 576–615
[42] Dirac P A M 1964 Lectures on Quantum Mechanics (New York: Belfer Graduate School of Sciences, Yeshiva University)
[43] Miković A, Oliveira M A and Vojinović M 2016 Hamiltonian analysis of the BFCG theory for the Poincaré 2-group Class. Quantum Grav. 33 065007
[49] Celada M and Montesinos M 2012 Lorentz-covariant Hamiltonian analysis of BF gravity with the Immirzi parameter Class. Quantum Grav. 29 205010
[50] Escalante A and Rubalcava-García I 2012 A pure Dirac’s canonical analysis for four-dimensional BF theories Int. J. Geom. Methods Mod. Phys. 09 1250053
[51] Castellani L 1982 Symmetries in constrained Hamiltonian systems Ann. Phys. 143 357
[52] Kichenassamy S 1986 Lagrange multipliers in theories of gravitation Ann. Phys. 168 404–24
[53] Ferraris M, Francaviglia M and Reina C 1982 Variational formulation of general relativity from 1915 to 1925 ‘Palatini’s method’ discovered by Einstein in 1925 Gen. Relativ. Gravit. 14 243
[54] Alexandrov S 2006 Reality conditions for Ashtekar gravity from Lorentz-covariant formulation Class. Quantum Grav. 23 1837–50
[55] Charles C 2018 Simplicity constraints: a 3D toy-model for loop quantum gravity Phys. Rev. D 97 1060002
[56] Caselle M, D’Adda A and Magnea L 1989 Regge calculus as a local theory of the Poincaré group Phys. Lett. B 232 457
[57] Gionti G 2005 Discrete gravity as a local theory of the Poincaré group in the first-order formalism Class. Quantum Grav. 22 4217–30