INSTITUTE OF PHYSICS PUBLISHING CLASSICAL AND QUANTUM GRAVITY

Class. Quantum Grav. **20** (2003) R43–R104 PII: S0264-9381(03)31513-8

TOPICAL REVIEW

Spin foam models for quantum gravity

Alejandro Perez

Center for Gravitational Physics and Geometry, The Pennsylvania State University, University Park, PA 16802, USA
and
Erwin Schrödinger International Institute for Mathematical Physics, Boltzmanngasse 9, A-1090, Wien, Austria

Received 4 September 2002, in final form 10 December 2002
Published 21 February 2003
Online at stacks.iop.org/CQG/20/R43

Abstract

In this topical review, we review the present status of the spin foam formulation of non-perturbative (background-independent) quantum gravity. The topical review is divided into two parts. In the first part, we present a general introduction to the main ideas emphasizing their motivation from various perspectives. Riemannian three-dimensional gravity is used as a simple example to illustrate conceptual issues and the main goals of the approach. The main features of the various existing models for four-dimensional gravity are also presented here. We conclude with a discussion of important questions to be addressed in four dimensions (gauge invariance, discretization independence, etc).

In the second part, we concentrate on the definition of the Barrett–Crane model. We present the main results obtained in this framework from a critical perspective. Finally, we review the combinatorial formulation of spin foam models based on the dual group field theory technology. We present the Barrett–Crane model in this framework and review the finiteness results obtained for both its Riemannian and its Lorentzian variants.

PACS number: 04.60.Pp

Contents

1. Introduction

Quantum gravity, the theory expected to reconcile the principles of quantum mechanics and general relativity, remains a major challenge in theoretical physics (for a review of the history of quantum gravity, see [1]). The main lesson of general relativity is that, unlike in any other interaction, spacetime geometry is fully dynamical. This special feature of gravity precludes the possibility of representing fields on a fixed background geometry and severely constrains the applicability of standard techniques that are successful in the description of other interactions. Although the necessity of a background-independent formulation of quantum gravity is widely recognized, there is a current debate about the means by which it should be implemented. In particular, it is not clear whether the non-renormalizability of perturbative quantum gravity should be interpreted as an indication of the inconsistency of general relativity at high energies, the inconsistency of the background-dependent framework applied to gravity or a combination of both.

According to the (background-dependent) perspective of standard QFT [2], non-renormalizability signals the inconsistency of the theory at high energies to be corrected by a more fundamental theory in the UV regime. A classical example of this is Fermi's four-fermion theory as an effective description of the weak interaction. According to this view different approaches to quantum gravity have been defined in terms of modifications of general relativity based on supersymmetry, higher dimensions, strings, etc. The finiteness properties of the perturbative expansions (which are background dependent from the onset) are improved in these theories; however, the definition of a background-independent quantization of such modifications remains open.

The approach of non-perturbative quantum gravity is based on a different interpretation of the infinities in perturbative quantum gravity: it is precisely the perturbative (background-dependent) techniques which are inconsistent with the fundamental nature of gravity. This view is strongly suggested by the predictions of the background-independent canonical quantization of general relativity defined by loop quantum gravity. Loop quantum gravity (LQG) is a non-perturbative formulation of quantum gravity based on the connection formulation of general relativity (for reviews on the subject, see [3–6]). A great deal of progress has been made within the theory. At the mathematical level, the main achievement is the rigorous definition of the Hilbert space of quantum geometry, the regularization of geometric operators and the rigorous definition of the quantum Hamiltonian constraint (defining the quantum dynamics). States of quantum geometry are given by polymer-like excitations supported on graphs (spin network states). From the physical viewpoint its main prediction is the discreteness of geometry at the Planck scale. This provides a clear-cut understanding of the problem of UV divergences in perturbative general relativity: at the Planck scale the classical notion of space and time simply ceases to exist; therefore, it is the assumption of a fixed smooth background geometry (typically flat spacetime) in perturbation theory that becomes inconsistent at high energies. The theory successfully incorporates interactions between quantum geometry and quantum matter in a way that is completely free of divergences [7]. The quantum nature of space appears as a physical regulator for the other interactions.

The dynamics is governed by the quantum Hamiltonian constraint. Even when this operator is rigorously defined [8] it is technically difficult to characterize its solution space. This is partly because the 3 + 1 decomposition of spacetime (necessary in the canonical formulation) breaks the manifest 4-diffeomorphism invariance of the theory making awkward the analysis of dynamics. The situation is somewhat analogous to that in standard quantum field theory. In the Hamiltonian formulation of standard quantum field theory manifest Lorentz invariance is lost due to a particular choice of time slicing of Minkowski spacetime. The formalism is certainly Lorentz invariant, but one has to work harder to show it explicitly. Manifest Lorentz invariance can be kept only in the Lagrangian (path integral) quantization making the (formal) path integral a powerful device for analysing relativistic dynamics.

Consequently, there has been growing interest in trying to define the dynamics in loop quantum gravity from a four-dimensional covariant perspective. This has given rise to the so-called spin foam approach to quantum gravity. Its main idea is the construction of a rigorous definition of the path integral for gravity based on the deep insights obtained in the canonical framework of loop quantum gravity. In turn, the path integral provides a device to explicitly solve the dynamics: path integral transition amplitudes can be shown to correspond to solutions of the quantum Hamiltonian constraint.

The underlying discreteness discovered in loop quantum gravity is crucial: in spin foam models the formal Misner–Hawking functional integral for gravity is replaced by a sum over combinatorial objects given by foam-like configurations (spin foams). A spin foam represents a possible history of the gravitational field and can be interpreted as a set of transitions through different quantum states of space. Boundary data in the path integral are given by the polymer-like excitations (spin network states) representing 3-geometry states in loop quantum gravity. General covariance implies the absence of a meaningful notion of time, and transition amplitudes are to be interpreted as defining the physical scalar product.

While the construction can be explicitly carried out in three dimensions there are additional technical difficulties in four dimensions. Various models have been proposed. A natural question is whether the infinite sums over geometries defining transition amplitudes would converge. In fact, there is no UV problem due to the fundamental discreteness and potential divergences are associated with the IR regime. There are recent results in the context of the Barrett–Crane model showing that amplitudes are well defined when the topology of the histories is restricted in a certain way.

The aim of this topical review is to provide a comprehensive review of the progress that has been achieved in the spin foam approach over the last few years and provide as well a self-contained introduction for the interested reader who is not familiar with the subject. The topical review is divided into two fundamental parts. In the first part, we present a general introduction to the subject including a brief summary of LQG in section 2. We introduce the spin foam formulation from different perspectives in section 3. In section 4, we present a simple example of the spin foam model: Riemannian three-dimensional gravity. We use this example as the basic tool to introduce the main ideas and to illustrate various conceptual issues. We review the different proposed models for four-dimensional quantum gravity in section 5. Finally, in section 6, we conclude the first part by analysing the various conceptual issues that arise in the approach. The first part is a general introduction to the formalism; it is self-contained and could be read independently.

One of the simplest and most studied spin foam models for four-dimensional gravity is the Barrett–Crane model [9, 10]. The main purpose of the second part is to present a critical survey of the different results that have been obtained in this framework and its combinatorial generalizations [11–13] based on the dual group field theory (GFT) formulation. In section 7, we present a systematic derivation of the Barrett–Crane model from the *Spin*(4) Plebanski's formulation. This derivation follows an alternative path from that of [14, 15]; here we emphasize the connection to a simplicial action.

Spin foams can be thought of as Feynman diagrams. In fact a wide class of spin foam models can be derived from the perturbative (Feynman) expansion of certain dual group field theories [16, 17]. A brief review of the main ideas involved is presented in section 8. We conclude the second part by studying the GFT formulation of the Barrett–Crane model for both Riemannian and Lorentzian geometry. The definition of the actual models and the sketch of the corresponding finiteness proofs [18–20] are given in section 9.

Part I: Main ideas

2. Loop quantum gravity and quantum geometry

Loop quantum gravity is a rigorous realization of the quantization programme established in the 1960s by Dirac, Wheeler, DeWitt, among others (for recent reviews, see [3, 4, 6]).

The technical difficulties of Wheeler's 'geometrodynamics' are circumvented by the use of connection variables instead of metrics [21–23]. At the kinematical level, the formulation is similar to that of standard gauge theories. The fundamental difference is, however, the absence of any non-dynamical background field in the theory.

The configuration variable is an $SU(2)$-connection A_a^i on a 3-manifold Σ representing space. The canonical momenta are given by the densitized triad E_i^a. The latter encode the (fully dynamical) Riemannian geometry of Σ and are the analogue of the 'electric fields' of Yang–Mills theory.

In addition to diffeomorphisms there is the local $SU(2)$ gauge freedom that rotates the triad and transforms the connection in the usual way. According to Dirac, gauge freedoms result in constraints among the phase space variables which conversely are the generating functionals of infinitesimal gauge transformations. In terms of connection variables the constraints are

$$\mathcal{G}_i = \mathcal{D}_a E_i^a = 0, \qquad \mathcal{C}_a = E_k^b F_{ba}^k = 0, \qquad \mathcal{S} = \epsilon^{ijk} E_i^a E_j^b F_{abk} + \cdots = 0, \tag{1}$$

where $\mathcal{D}_a$ is the covariant derivative and F_{ba} is the curvature of A_a^i. $\mathcal{G}_i$ is the familiar Gauss constraint—analogous to the Gauss law of electromagnetism—generating infinitesimal $SU(2)$ gauge transformations, $\mathcal{C}_a$ is the vector constraint generating space diffeomorphism and $\mathcal{S}$ is the scalar constraint generating 'time' reparametrization (there is an additional term that we have omitted for simplicity).

Loop quantum gravity is defined using Dirac quantization. One first represents (1) as operators in an auxiliary Hilbert space $\mathcal{H}$ and then solves the constraint equations

$$\hat{\mathcal{G}}_i \Psi = 0, \qquad \hat{\mathcal{C}}_a \Psi = 0, \qquad \hat{\mathcal{S}} \Psi = 0. \tag{2}$$

The Hilbert space of solutions is the so-called physical Hilbert space $\mathcal{H}_{\text{phys}}$. In a generally covariant system quantum dynamics is fully governed by constraint equations. In the case of loop quantum gravity they represent *quantum Einstein's equations*.

States in the auxiliary Hilbert space are represented by wave functionals of the connection $\Psi(A)$ which are square integrable with respect to a natural diffeomorphism-invariant measure, the Ashtekar–Lewandowski measure [24] (we denote it by $\mathcal{L}^2[\mathcal{A}]$, where $\mathcal{A}$ is the space of (generalized) connections). This space can be decomposed into a direct sum of orthogonal subspaces $\mathcal{H} = \bigoplus_\gamma \mathcal{H}_\gamma$ labelled by a graph γ in Σ. The fundamental excitations are given by the holonomy $h_\ell(A) \in SU(2)$ along a path ℓ in Σ:

$$h_\ell(A) = \mathcal{P} \exp \int_\ell A. \tag{3}$$

Elements of $\mathcal{H}_\gamma$ are given by functions

$$\Psi_{f,\gamma}(A) = f\left(h_{\ell_1}(A), h_{\ell_2}(A), \ldots, h_{\ell_n}(A)\right), \tag{4}$$

where h_ℓ is the holonomy along the links $\ell \in \gamma$ and $f : SU(2)^n \to \mathbb{C}$ is (Haar measure) square integrable. They are called *cylindrical functions* and represent a dense set in $\mathcal{H}$ denoted by Cyl.

Gauge transformations generated by the Gauss constraint act non-trivially at the endpoints of the holonomy, i.e., at the nodes of graphs. The Gauss constraint (in (1)) is solved by looking at $SU(2)$ gauge-invariant functionals of the connection ($\mathcal{L}^2[\mathcal{A}]/\mathcal{G}$). The fundamental gauge-invariant quantity is given by the holonomy around closed loops. An orthonormal basis of the kernel of the Gauss constraint is defined by the so-called spin network states $\Psi_{\gamma,\{j_\ell\},\{\iota_n\}}(A)$ [25–27]. Spin networks[1] are defined by a graph γ in Σ, a collection of spins $\{j_\ell\}$—unitary irreducible representations of $SU(2)$—associated with links $\ell \in \gamma$ and a collection

[1] Spin networks were introduced by Penrose [28] in a attempt to define three-dimensional Euclidean quantum geometry from the combinatorics of angular momentum in QM. Independently they have been used in lattice gauge theory [29, 30] as a natural basis for gauge-invariant functions on the lattice. For an account of their applications in various contexts, see [31].

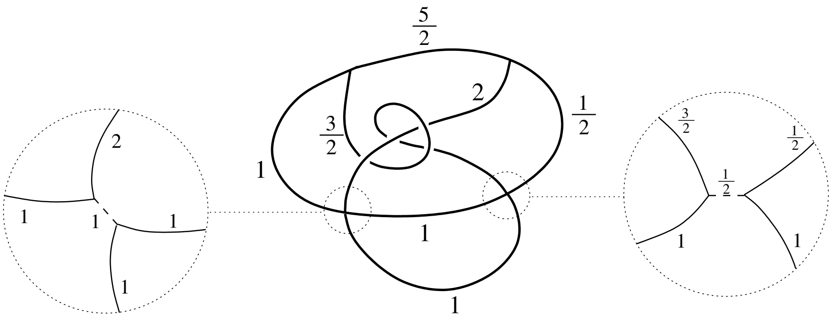


Figure 1. Spin network state: at three-valent nodes the intertwiner is uniquely specified by the corresponding spins. At four or higher valent nodes an intertwiner has to be specified. Choosing an intertwiner corresponds to decomposing the *n*-valent node in terms of three-valent ones adding new virtual links (dashed lines) and their corresponding spins. This is illustrated explicitly in the figure for the two four-valent nodes.

of $SU(2)$ intertwiners $\{\iota_n\}$ associated with nodes $n \in \gamma$ (see figure 1). The spin network gauge-invariant wave functional $\Psi_{\gamma,\{j_\ell\},\{\iota_n\}}(A)$ is constructed by first associating an $SU(2)$ matrix in the j_ℓ-representation with the holonomies $h_\ell(A)$ corresponding to the link ℓ, and then contracting the representation matrices at nodes with the corresponding intertwiners ι_n, namely

$$\Psi_{\gamma,\{j_\ell\},\{\iota_n\}}(A) = \prod_{n\in\gamma} \iota_n \prod_{\ell\in\gamma} j_\ell[h_\ell(A)], \tag{5}$$

where $j_\ell[h_\ell(A)]$ denotes the corresponding j_ℓ-representation matrix evaluated at the corresponding link holonomy and the matrix index contraction is left implicit.

The solution of the vector constraint is more subtle [24]. One uses group averaging techniques and the existence of the diffeomorphism-invariant measure. The fact that zero lies in the continuous spectrum of the diffeomorphism constraint implies solutions to correspond to generalized states. These are not in $\mathcal{H}$ but are elements of the topological dual Cyl^*.[2] However, the intuitive idea is quite simple: solutions to the vector constraint are given by equivalence classes of spin network states up to diffeomorphism. Two spin network states are considered equivalent if their underlying graphs can be deformed into each other by the action of a diffeomorphism.

This can be regarded as an indication that the smooth spin network category could be replaced by something which is more combinatorial in nature so that diffeomorphism invariance becomes a derived property of the classical limit. LQG has been modified along these lines by replacing the smooth manifold structure of the standard theory by the weaker concept of piecewise linear manifold [32]. In this context, graphs defining spin network states can be completely characterized using the combinatorics of cellular decompositions of space. Only a discrete analogue of the diffeomorphism symmetry survives which can be dealt with in a fully combinatorial manner. We will take this point of view when we introduce the notion of spin foam in the following section.

2.1. Quantum geometry

The generalized states described above solve all the constraints (1) but the scalar constraint. They are regarded as quantum states of the Riemannian geometry on Σ. They define the kinematical sector of the theory known as *quantum geometry*.

[2] According to the triple $Cyl \subset \mathcal{H} \subset Cyl^*$.

Geometric operators acting on spin network states can be defined in terms of the fundamental triad operators $\hat{E}^a_i$. The simplest of such operators is the area of a surface S classically given by

$$A_S(E) = \int_S \mathrm{d}x^2 \sqrt{\mathrm{Tr}[n_a n_b E^a E^b]}, \tag{6}$$

where n is a conormal. The geometric operator $\hat{A}_S(E)$ can be rigorously defined by its action on spin network states [33–35]. The area operator gives a clear geometrical interpretation to spin network states: the fundamental one-dimensional excitations defining a spin network state can be thought of as quantized 'flux lines' of area. More precisely, if the surface $S \subset \Sigma$ is punctured by a spin network link carrying a spin j, this state is an eigenstate of $\hat{A}_S(E)$ with eigenvalue proportional to $\ell_P^2\sqrt{j(j+1)}$. In the generic sector, where no node lies on the surface, the spectrum takes the simple form

$$a_S(\{j\}) = 8\pi\iota\ell_P^2 \sum_i \sqrt{j_i(j_i+1)}, \tag{7}$$

where i labels punctures and ι is the Immirzi parameter [36].[3] $a_S(\{j\})$ is the sum of single puncture contributions. The general form of the spectrum including the cases where nodes lie on S has been computed in closed form [35].

The spectrum of the volume operator is also discrete [33, 34, 38, 39]. If we define the volume operator $\hat{V}_\sigma(E)$ of a three-dimensional region $\sigma \subset \Sigma$ then non-vanishing eigenstates are given by spin networks containing n-valent nodes in σ for $n > 3$. Volume is concentrated in nodes.

2.2. *Quantum dynamics*

In contrast to the Gauss and vector constraints, the scalar constraint does not have a simple geometrical meaning. This makes its quantization more involved. Regularization choices have to be made and the result is not unique. After Thiemann's first rigorous quantization [40] other well-defined possibilities have been found [41–43]. This ambiguity affects the dynamics governed by

$$\hat{\mathcal{S}}\Psi = 0. \tag{8}$$

The difficulty in dealing with the scalar constraint is not surprising. The vector constraint, generating space diffeomorphisms, and the scalar constraint, generating time reparametrizations, arise from the underlying 4-diffeomorphism invariance of gravity. In the canonical formulation the 3 + 1 splitting breaks the manifest four-dimensional symmetry. The price paid is the complexity of the time reparametrization constraint $\mathcal{S}$. The situation is somewhat reminiscent of that in standard quantum field theory where manifest Lorentz invariance is lost in the Hamiltonian formulation[4].

From this perspective, there has been growing interest in approaching the problem of dynamics by defining a covariant formulation of quantum gravity. The idea is that (as in the QFT case) one can keep manifest four-dimensional covariance in the path integral formulation. The spin foam approach is an attempt to define the path integral quantization of gravity using what we have learned from LQG.

[3] The Immirzi parameter ι is a free parameter in the theory. This ambiguity is purely quantum mechanical (it disappears in the classical limit). It has to be fixed in terms of physical predictions. The computation of black-hole entropy in LQG fixes the value of ι (see [37]).

[4] There is, however, an additional complication here: the canonical constraint algebra does not reproduce the 4-diffeomorphism Lie algebra. This complicates the geometrical meaning of S.

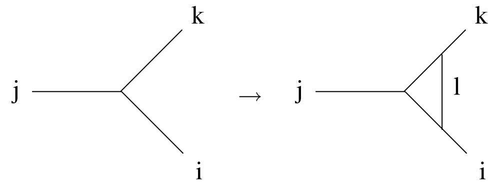


Figure 2. A typical transition generated by the action of the scalar constraint.

In standard quantum mechanics path integrals provide the solution of dynamics as a device to compute the time evolution operator. Similarly, in the generally covariant context it provides a tool to find solutions to the constraint equations (this has been emphasized formally in various places: in the case of gravity see, for example, [44], for a detailed discussion of this in the context of quantum mechanics, see [45]). We will come back to this issue later.

Let us conclude by stating some properties of $\hat{S}$ that do not depend on the ambiguities mentioned above. One is the discovery that smooth loop states naturally solve the scalar constraint operator [46, 47]. This set of states is clearly too small to represent the physical Hilbert space (e.g., they span a zero volume sector). However, this implies that $\hat{S}$ acts only on spin network nodes. Its action modifies spin networks at nodes by creating new links according to figure 2.[5] This is crucial in the construction of the spin foam approach of the following section.

3. Spin foams and the path integral for gravity

The possibility of defining quantum gravity using Feynman's path integral approach was considered long ago by Misner and later extensively studied by Hawking, Hartle and others [48, 49]. Given a 4-manifold $\mathcal{M}$ with boundaries Σ_1 and Σ_2, and denoting by G the space of metrics on $\mathcal{M}$, the transition amplitude between $|[q_{ab}]\rangle$ on Σ_1 and $|[q'_{ab}]\rangle$ on Σ_2 is formally

$$\langle [q_{ab}] | [q'_{ab}] \rangle = \int_{[g]} \mathcal{D}[g]\, \mathrm{e}^{\mathrm{i}S([g])}, \tag{9}$$

where the integration on the right-hand side is performed over all spacetime metrics up to 4-diffeomorphisms $[g] \in G/_{Diff(\mathcal{M})}$ with fixed boundary values up to 3-diffeomorphisms $[q_{ab}]$, $[q'_{ab}]$, respectively.

There are various difficulties associated with (9). Technically there is the problem of defining the functional integration over $[g]$ on the RHS. This is partially because of the difficulties in defining infinite-dimensional functional integration beyond the perturbative framework. In addition, there is the issue of having to deal with the space $G/_{Diff(\mathcal{M})}$, i.e., how to characterize the diffeomorphism-invariant information in the metric. This gauge problem (3-diffeomorphisms) is also present in the definition of the boundary data. There is no well-defined notion of kinematical state $|[q_{ab}]\rangle$ as the notion of kinematical Hilbert space in standard metric variables has never been defined.

We can be more optimistic in the framework of loop quantum gravity. The notion of quantum state of 3-geometry is rigorously defined in terms of spin network states. They carry the diff-invariant information of the Riemannian structure of Σ. In addition, and very importantly, these states are intrinsically discrete (coloured graphs on Σ) suggesting

[5] This is not the case in all the available definitions of the scalar constraints, such as for example the one defined in [42, 43].

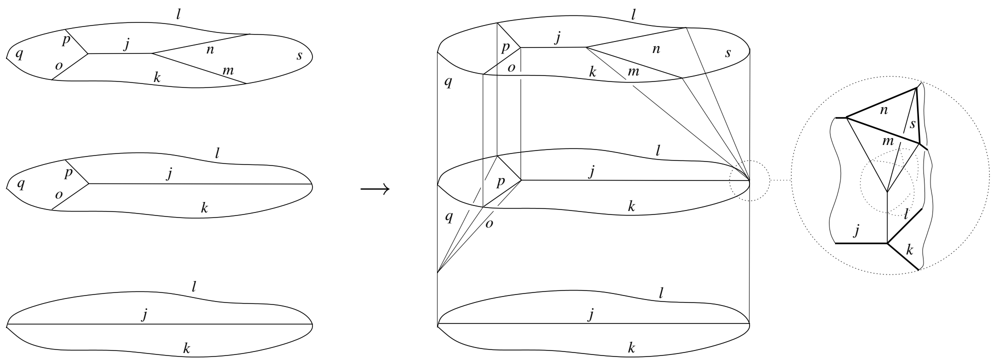


Figure 3. A typical path in a path integral version of loop quantum gravity is given by a series of transitions through different spin network states representing a state of 3-geometries. Nodes and links in the spin network evolve into one-dimensional edges and faces. New links are created and spins are reassigned at vertices (emphasized on the right). The 'topological' structure is provided by the underlying 2-complex while the geometric degrees of freedom are encoded in the labelling of its elements with irreducible representations and intertwiners.

a possible solution to the functional measure problem, i.e., the possibility of constructing a notion of Feynman 'path integral' in a combinatorial manner involving sums over spin network worldsheet amplitudes. Heuristically, '4-geometries' are to be represented by 'histories' of quantum states of 3-geometries or spin network states. These 'histories' involve a series of transitions between spin network states (figure 3), and define a foam-like structure (a '2-graph' or 2-complex) whose components inherit the spin representations from the underlying spin networks. These spin network worldsheets are the so-called *spin foams*.

The precise definition of spin foams was introduced by Baez in [14] emphasizing their role as morphisms in the category defined by spin networks[6]. A spin foam $\mathcal{F} : s \to s'$, representing a transition from the spin network $s = (\gamma, \{j_\ell\}, \{\iota_n\})$ into $s' = (\gamma', \{j_{\ell'}\}, \{\iota_{n'}\})$, is defined by a 2-complex[7] $\mathcal{J}$ bordered by the graphs of γ and γ', respectively, a collection of spins $\{j_f\}$ associated with faces $f \in \mathcal{J}$ and a collection of intertwiners $\{\iota_e\}$ associated with edges $e \in \mathcal{J}$. Both spins and intertwiners of exterior faces and edges match the boundary values defined by the spin networks s and s', respectively. Spin foams $\mathcal{F} : s \to s'$ and $\mathcal{F}' : s' \to s''$ can be composed into $\mathcal{F}\mathcal{F}' : s \to s''$ by gluing together the two corresponding 2-complexes at s'. A spin foam model is an assignment of amplitudes $A[\mathcal{F}]$ which is consistent with this composition rule in the sense that

$$A[\mathcal{F}\mathcal{F}'] = A[\mathcal{F}]A[\mathcal{F}']. \tag{10}$$

Transition amplitudes between spin network states are defined by

$$\langle s, s'\rangle_{\mathrm{phys}} = \sum_{\mathcal{F}:s\to s'} A[\mathcal{F}], \tag{11}$$

where the notation anticipates the interpretation of such amplitudes as defining the physical scalar product. The domain of the previous sum is left unspecified at this stage. We shall discuss this question further in section 6. This last equation is the spin foam counterpart of equation (9). This definition remains formal until we specify what the set of allowed spin foams in the sum is and define the corresponding amplitudes.

[6] The role of category theory for quantum gravity had been emphasized by Crane in [50–52].

[7] In most of this topical review we use the concept of piecewise linear 2-complexes as in [14]; in section 8, we shall study a formulation of spin foam in terms of certain combinatorial 2-complexes.

In standard quantum mechanics the path integral is used to compute the matrix elements of the evolution operator $U(t)$. It provides in this way the solution for dynamics since for any kinematical state Ψ the state $U(t)\Psi$ is a solution to Schrödinger's equation. Analogously, in a generally covariant theory the path integral provides a device for constructing solutions to the quantum constraints. Transition amplitudes represent the matrix elements of the so-called generalized 'projection' operator P (sections 3.1 and 6.3) such that $P\Psi$ is a physical state for any kinematical state Ψ. As in the case of the vector constraint the solutions of the scalar constraint correspond to distributional states (zero is in the continuum part of its spectrum). Therefore, $\mathcal{H}_{\rm phys}$ is not a proper subspace of $\mathcal{H}$ and the operator P is not a projector (P^2 is ill-defined)[8]. In section 4, we give an explicit example of this construction.

The background-independent character of spin foams is manifest. The 2-complex can be thought of as representing 'spacetime' while the boundary graphs represent 'space'. They do not carry any geometrical information in contrast to the standard concept of a lattice. Geometry is encoded in the spin labellings which represent the degrees of freedom of the gravitational field.

3.1. Spin foams and the projection operator into $\mathcal{H}_{\rm phys}$

Spin foams naturally arise in the formal definition of the exponentiation of the scalar constraint as studied by Reisenberger and Rovelli [53] and Rovelli [54]. The basic idea consists of constructing the 'projection' operator P providing a definition of the formal expression

$$P = \prod_{x\in\Sigma} \delta(\hat{\mathcal{S}}(x)) = \int \mathcal{D}[N]\,{\rm e}^{{\rm i}\hat{\mathcal{S}}[N]}, \tag{12}$$

where $\hat{\mathcal{S}}[N] = \int {\rm d}x^3 N(x)\hat{\mathcal{S}}(x)$, with $N(x)$ being the lapse function. P defines the *physical* scalar product according to

$$\langle s, s'\rangle_{\rm phys} = \langle sP, s'\rangle, \tag{13}$$

where the RHS is defined using the kinematical scalar product. Reisenberger and Rovelli make progress towards a definition of (12) by constructing a truncated version P_Λ (where Λ can be regarded as an infrared cut-off). One of the main ingredients is Rovelli's definition of a diffeomorphism-invariant measure $\mathcal{D}[N]$ generalizing the techniques of [55].

The starting point is the expansion of the exponential in (12) in powers

$$\langle sP_\Lambda, s'\rangle = \int_{|N(x)|\leqslant\Lambda} \mathcal{D}[N] \left\langle s \sum_{n=0}^{\infty} \frac{i^n}{n!}(\mathcal{S}[N])^n, s'\right\rangle. \tag{14}$$

The construction works for a generic form of quantum scalar constraint as long as it acts locally on spin network nodes both creating and destroying links (this local action generates a vertex of the type emphasized in figure 2). The action of $\hat{\mathcal{S}}[N]$ depends on the value of the lapse at nodes. Integration over the lapse can be performed and the final result is given by a power series in the cut-off Λ, namely

$$\langle sP_\Lambda, s'\rangle = \sum_{n=0}^{\infty} \frac{i^n\Lambda^n}{n!} \sum_{\mathcal{F}_n:s\to s'} A[\mathcal{F}_n] = \sum_{n=0}^{\infty} \frac{i^n\Lambda^n}{n!} \sum_{\mathcal{F}_n:s\to s'} \prod_v A_v(\rho_v, \iota_v), \tag{15}$$

where $\mathcal{F}_n : s \to s'$ are spin foams generated by n actions of the scalar constraint, i.e., spin foams with n vertices. The spin foam amplitude $A[\mathcal{F}_n]$ factorizes in a product of vertex contributions $A_v(\rho_v, \iota_v)$ depending on the spins j_v and ι_v neighbouring faces and edges.

[8] In the notation of the previous section states in $\mathcal{H}_{\rm phys}$ are elements of Cyl^*.

The spin foam shown in figure 3 corresponds in this context to two actions of $\mathcal{S}$ and would contribute to the amplitude in the order Λ^2.

Physical observables can be constructed out of kinematical operators using P. If $O_{\rm kin}$ represents an operator commuting with all but the scalar constraint then $O_{\rm phys} = P O_{\rm kin} P$ defines a physical observable. Its expectation value is

$$\langle O_{\rm phys}\rangle = \frac{\langle s P O_{\rm kin} P, s\rangle}{\langle s P, s\rangle} := \lim_{\Lambda\to\infty} \frac{\langle s P_\Lambda O_{\rm kin} P_\Lambda, s\rangle}{\langle s P_\Lambda, s\rangle}, \tag{16}$$

where the limit of the ratio of truncated quantities is expected to converge for suitable operators $O_{\rm kin}$. Issues of convergence have not been studied and they would clearly be regularization dependent.

3.2. Spin foams from lattice gravity

Spin foam models naturally arise in lattice discretizations of the path integral of gravity and generally covariant gauge theories. This was originally studied by Reisenberger [25]. The spacetime manifold is replaced by a lattice given by a cellular complex. The discretization allows for the definition of the functional measure reducing the number of degrees of freedom to finitely many. The formulation is similar to that of standard lattice gauge theory. However, the nature of this truncation is fundamentally different: background independence implies that it cannot be simply interpreted as a UV regulator (we will be more explicit in the following).

We present a brief outline of the formulation (for details, see [25, 56, 57]). Start from the action of gravity in some first-order formulation ($S(e, A)$). The formal path integral takes the form

$$Z = \int \mathcal{D}[e]\mathcal{D}[A]\, \mathrm{e}^{\mathrm{i}S(e,A)} = \int \mathcal{D}[A]\, \mathrm{e}^{\mathrm{i}S_{\rm eff}(A)}, \tag{17}$$

where in the second part we have formally integrated over the tetrad e obtaining an effective action $S_{\rm eff}$.[9] From this point on, the derivation is analogous to that of generally covariant gauge theory. The next step is to define the previous equation on a 'lattice'.

As for Wilson's action for standard lattice gauge theory the relevant structure for the discretization is a 2-complex $\mathcal{J}$. We assume the 2-complex to be defined in terms of the dual 2-skeleton $\mathcal{J}_\Delta$ of a simplicial complex Δ. Denoting the edges by $e \in \mathcal{J}_\Delta$ and the plaquettes or faces by $f \in \mathcal{J}_\Delta$, one discretizes the connection by assigning a group element g_e to edges

$$A \to \{g_e\}.$$

The Haar measure on the group is used to represent the connection integration:

$$\mathcal{D}[A] \to \prod_{e\in\mathcal{J}_\Delta} \mathrm{d}g_e.$$

The action of gravity depends on the connection A only through the curvature $F(A)$ so that upon discretization the action is expressed as a function of the holonomy around faces g_f corresponding to the product of the g_e which we denote by $g_f = g_e^{(1)} \cdots g_e^{(n)}$:

$$F(A) \to \{g_f\}.$$

In this way, $S_{\rm eff}(A) \to S_{\rm eff}(\{g_f\})$. Thus the lattice path integral becomes

$$Z = \int \prod_{e\in\mathcal{J}_\Delta} \mathrm{d}g_e \exp[\mathrm{i}S_{\rm eff}(\{g_f\})]. \tag{18}$$

[9] This is a simplifying assumption in the derivation. One could put the full action in the lattice [56] and then integrate over the discrete e to obtain the discretized version of the quantity on the right-hand side of (17). This is what we will do in section 4.

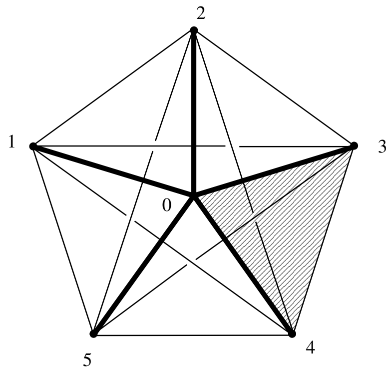


Figure 4. A fundamental *atom* is defined by the intersection of a dual vertex in $\mathcal{J}_\Delta$ (corresponding to a 4-simplex in Δ) with a 3-sphere. The thick lines represent the internal edges while the thin lines the intersections of the internal faces with the boundary. They define the boundary graph denoted by γ_5 below. One of the faces has been emphasized.

Reisenberger assumes that $S_{\text{eff}}(\{g_f\})$ is local in the following sense: the amplitude of any piece of the 2-complex obtained as its intersection with a ball depends only on the value of the connection on the corresponding boundary. The degrees of freedom communicate through the lattice connection on the boundary. One can compute the amplitudes of pieces of $\mathcal{J}_\Delta$ (at fixed boundary data) and then obtain the full $\mathcal{J}_\Delta$ amplitude by gluing the pieces together and integrating out the mutual boundary connections along common boundaries. The boundary of a portion of $\mathcal{J}_\Delta$ is a graph. The boundary value is an assignment of group elements to its links. The amplitude is a function of the boundary connection, i.e., an element of Cyl. In the case of a cellular 2-complex there is a maximal splitting corresponding to cutting out a neighbourhood around each vertex. If the discretization is based on the dual of a triangulation these elementary building blocks are all alike and denoted *atoms*. Such an atom in four dimensions is represented in figure 4.

The atom amplitude depends on the boundary data given by the value of the holonomies on the ten links of the pentagonal boundary graph γ_5 shown in the figure. This amplitude can be represented by a function

$$\mathcal{V}(\alpha_{ij}) \qquad \text{for} \quad \alpha_{ij} \in G \quad \text{and} \quad i \neq j = 1, \ldots, 5, \tag{19}$$

where α_{ij} represents the boundary lattice connection along the link ij in figure 4. Gauge invariance $\left(V(\alpha_{ij}) = V\left(g_i \alpha_{ij} g_j^{-1}\right)\right)$ implies that the function can be spanned in terms of spin network functions $\Psi_{\gamma_5, \{\rho_{ij}\}, \{\iota_i\}}(\alpha_{ij})$ based on the pentagonal graph γ_5, namely

$$\mathcal{V}(\alpha_{ij}) = \sum_{\rho_{ij}} \sum_{\iota_i} \tilde{\mathcal{V}}(\{\rho_{ij}\}, \{\iota_i\}) \Psi_{\gamma_5, \{\rho_{ij}\}, \{\iota_i\}}(\alpha_{ij}), \tag{20}$$

where $\tilde{\mathcal{V}}(\{\rho_{ij}\}, \{\iota_i\})$ is the atom amplitude in 'momentum' space depending on ten spins ρ_{ij} labelling the faces and five intertwiners ι_i labelling the edges. Gluing the atoms together the integral over common boundaries is replaced by the sum over common values of spin labels and intertwiners[10]. The total amplitude becomes

$$Z[\mathcal{J}_\Delta] = \sum_{\mathcal{C}_f : \{f\} \to \{\rho_f\}} \sum_{\mathcal{C}_e : \{e\} \to \{\iota_e\}} \prod_{f \in \mathcal{J}_\Delta} A_f(\rho_f, \iota_f) \prod_{v \in \mathcal{J}_\Delta} \tilde{\mathcal{V}}(\{\rho_v\}, \{\iota_v\}), \tag{21}$$

[10] This is a consequence of the basic orthogonality of unitary representations, namely

$$\int j[g]_{\alpha\beta} k[g]_{\gamma\mu} \, \mathrm{d}g = \frac{\delta_{jk^*}}{2j+1} \delta_{\alpha\gamma} \delta_{\beta\mu}.$$

where $\mathcal{C}_e : \{e\} \to \{\iota_e\}$ denotes the assignment of intertwiners to edges, $\mathcal{C}_f : \{\rho_f\} \to \{f\}$ the assignment of spins ρ_f to faces and $A_f(\rho_f, \iota_f)$ is the face amplitude arising in the integration over the lattice connection. The lattice definition of the path integral for gravity and covariant gauge theories becomes a discrete sum of spin foam amplitudes!

3.3. Spin foams for gravity from BF theory

The integration over the tetrad we formally performed in (17) is not always possible in practice. There is, however, a type of generally covariant theory for which the analogue integration is trivial. This is the case of a class of theories called BF theory. General relativity in three dimensions is of this type. Consequently, BF theory can be quantized along the lines of the previous section. BF spin foam amplitudes are simply given by certain invariants in the representation theory of the gauge group. We shall study in some detail the case of three-dimensional Riemannian quantum gravity in the following section.

The relevance of BF theory is its close relation with general relativity in four dimensions. In fact, general relativity can be described by certain BF theory action plus Lagrange multiplier terms imposing certain algebraic constraints on the fields [58]. This is the starting point for the definition of several of the models we will present in this topical review: a spin foam model for gravity can be defined by imposing restrictions on the spin foams that enter in the partition function of the BF theory. These restrictions are essentially the translation into representation theory of the constraints that reduce BF theory to general relativity.

3.4. Spin foams as Feynman diagrams

As already pointed out in [14] spin foams can be interpreted in close analogy with Feynman diagrams. Standard Feynman graphs are generalized to 2-complexes and the labelling of propagators by momenta to the assignment of spins to faces. Finally, momentum conservation at vertices in standard Feynmanology is now represented by spin conservation at edges, ensured by the assignment of the corresponding intertwiners. In spin foam models the non-trivial content of amplitudes is contained in the vertex amplitude as pointed out in sections 3.1 and 3.2 which in the language of Feynman diagrams can be interpreted as an interaction. We shall see that this analogy is indeed realized in the formulation of spin foam models in terms of a group field theory (GFT) [16, 17].

4. Spin foams for three-dimensional gravity

Three-dimensional gravity is an example of BF theory for which the spin foam approach can be implemented in a rather simple way. Despite its simplicity the theory allows for the study of many of the conceptual issues to be addressed in four dimensions. In addition, as we mentioned in section 3.3, spin foams for BF theory are the basic building block of four-dimensional gravity models. For a beautiful presentation of BF theory and its relation to spin foams, see [59]. For simplicity we study the Riemannian theory; the Lorentzian generalization of the results of this section have been studied in [60].

4.1. The classical theory

Riemannian gravity in three dimensions is a theory with no local degrees of freedom, i.e., no gravitons. Its action is given by

$$S(e, A) = \int_{\mathcal{M}} \mathrm{Tr}(e \wedge F(A)), \qquad (22)$$

where the field A is an $SU(2)$ connection and the triad e is a Lie-algebra-valued 1-form. The local symmetries of the action are $SU(2)$ gauge transformations

$$\delta e = [e, \omega], \qquad \delta A = d_A \omega, \tag{23}$$

where ω is an $su(2)$-valued 0-form, and 'topological' gauge transformation

$$\delta e = d_A \eta, \qquad \delta A = 0, \tag{24}$$

where d_A denotes the covariant exterior derivative and η is an $su(2)$-valued 0-form. The first invariance is manifest from the form of the action, while the second is a consequence of the Bianchi identity, $d_A F = 0$. The gauge freedom is so large that the theory has only global degrees of freedom. This can be checked directly by writing the equations of motion

$$F(A) = d_A A = 0, \qquad d_A e = 0. \tag{25}$$

The first implies that the connection is flat which in turn means that it is locally gauge ($A = d_A \omega$). The solutions of the second equation are also locally gauge as any closed form is locally exact ($e = d_A \eta$).[11] This very simple theory can be quantized in a direct manner.

4.2. Canonical quantization

Assuming $\mathcal{M} = \Sigma \times \mathrm{R}$ where Σ is a Riemann surface representing space, the phase space of the theory is parametrized by the spatial connection A (for simplicity we use the same notation as for the spacetime connection) and its conjugate momentum E. The constraints that result from the gauge freedoms (23) and (24) are

$$\mathcal{D}_a E_i^a = 0, \qquad F(A) = 0. \tag{26}$$

The first is the familiar Gauss constraint (in the notation of equation (1)), and we call the second the *curvature constraint*. There are six independent constraints for the six components of the connection, i.e., no local degrees of freedom. The kinematical setting is analogous to that of four-dimensional gravity and the quantum theory is defined along similar lines. The kinematical Hilbert space (quantum geometry) is spanned by $SU(2)$ spin network states on Σ which automatically solve the Gauss constraint. The 'dynamics' is governed by the curvature constraint. The physical Hilbert space is obtained by restricting the connection to be flat and the physical scalar product is defined by a natural measure in the space of flat connections [59]. The distributional character of the solutions of the curvature constraint is manifest here. Different spin network states are physically equivalent when they differ by a null state (states with vanishing physical scalar product with all spin network states). This happens when the spin networks are related by certain skein relations. One can reconstruct $\mathcal{H}_{\text{phys}}$ directly from the skein relations which, in turn, can be found by studying the covariant spin foam formulation of the theory.

4.3. Spin foam quantization of 3D gravity

Here we apply the general framework of section 3.2. This has been studied by Iwasaki in [61, 62]. The partition function, $\mathcal{Z}$, is formally given by[12]

$$\mathcal{Z} = \int \mathcal{D}[e]\mathcal{D}[A]\, e^{i\int_{\mathcal{M}} \mathrm{Tr}[e\wedge F(A)]}, \tag{27}$$

[11] One can easily check that the infinitesimal diffeomorphism gauge action $\delta e = \mathcal{L}_v e$, and $\delta A = \mathcal{L}_v A$, where $\mathcal{L}_v$ is the Lie derivative in the v direction, is a combination of (23) and (24) for $\omega = v^a A_a$ and $\eta = v^a e_a$, respectively, acting on the space of solutions, i.e. when (25) holds.

[12] We are dealing with Riemannian three-dimensional gravity. This should not be confused with the approach of Euclidean quantum gravity formally obtained by a Wick rotation of Lorentzian gravity. Note the imaginary unit in front of the action. The theory of Riemannian quantum gravity should be regarded as a toy model with no obvious connection to the Lorentzian sector.

where for the moment we assume $\mathcal{M}$ to be compact and orientable. Integrating over the e field in (27) we obtain

$$\mathcal{Z} = \int \mathcal{D}[A]\delta(F(A)). \tag{28}$$

The partition function $\mathcal{Z}$ corresponds to the 'volume' of the space of flat connections on $\mathcal{M}$.

In order to give a meaning to the formal expressions above, we replace the three-dimensional manifold $\mathcal{M}$ with an arbitrary cellular decomposition Δ. We also need the notion of the associated dual 2-complex of Δ denoted by $\mathcal{J}_\Delta$. The dual 2-complex $\mathcal{J}_\Delta$ is a combinatorial object defined by a set of vertices $v \in \mathcal{J}_\Delta$ (dual to 3-cells in Δ), edges $e \in \mathcal{J}_\Delta$ (dual to 2-cells in Δ) and faces $f \in \mathcal{J}_\Delta$ (dual to 1-cells in Δ).

The fields e and A have support on these discrete structures. The $su(2)$-valued 1-form field e is represented by the assignment of an $e \in su(2)$ to each 1-cell in Δ. The connection field A is represented by the assignment of group elements $g_e \in SU(2)$ to each edge in $\mathcal{J}_\Delta$.

The partition function is defined by

$$\mathcal{Z}(\Delta) = \int \prod_{f \in \mathcal{J}_\Delta} \mathrm{d}e_f \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \, \mathrm{e}^{\mathrm{i}\mathrm{Tr}[e_f U_f]}, \tag{29}$$

where $\mathrm{d}e_f$ is the regular Lebesgue measure on R^3, $\mathrm{d}g_e$ is the Haar measure on $SU(2)$ and U_f denotes the holonomy around faces, i.e., $U_f = g_e^1 \cdots g_e^N$ for N being the number of edges bounding the corresponding face. Since $U_f \in SU(2)$ we can write it as $U_f = u_f^0 1 + F_f$, where $u_f^0 \in \mathbb{C}$ and $F_f \in su(2)$. F_f is interpreted as the discrete curvature around the face f. Clearly $\mathrm{Tr}[e_f U_f] = \mathrm{Tr}[e_f F_f]$. An arbitrary orientation is assigned to faces when computing U_f. We use the fact that faces in $\mathcal{J}_\Delta$ are in one-to-one correspondence with 1-cells in Δ and label e_f with a face subindex.

Integrating over e_f, we obtain

$$\mathcal{Z}(\Delta) = \int \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \prod_{f \in \mathcal{J}_\Delta} \delta\left(g_e^1 \cdots g_e^N\right), \tag{30}$$

where δ corresponds to the delta distribution defined on $\mathcal{L}^2(SU(2))$. Note that the previous equation corresponds to the discrete version of equation (28).

Integration over the discrete connection $\left(\prod_e \mathrm{d}g_e\right)$ can be performed by expanding first the delta function in the previous equation using the Peter–Weyl decomposition [63]

$$\delta(g) = \sum_{j \in \mathrm{irrep}(SU(2))} \Delta_j \, \mathrm{Tr}[j(g)], \tag{31}$$

where $\Delta_j = 2j + 1$ denotes the dimension of the unitary representation j, and $j(g)$ is the corresponding representation matrix. Using equation (31), the partition function (30) becomes

$$\mathcal{Z}(\Delta) = \sum_{\mathcal{C}:\{j\} \to \{f\}} \int \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \prod_{f \in \mathcal{J}_\Delta} \Delta_{j_f} \mathrm{Tr}\left[j_f\left(g_e^1 \cdots g_e^N\right)\right], \tag{32}$$

where the sum is over colouring of faces in the notation of (21).

Going from equation (29) to (32) we have replaced the continuous integration over the e by the sum over representations of $SU(2)$. Roughly speaking, the degrees of freedom of e are now encoded in the representation being summed over in (32).

Now it remains to integrate over the lattice connection $\{g_e\}$. If an edge $e \in \mathcal{J}_\Delta$ bounds n faces there are n traces of the form $\mathrm{Tr}[j_f(\cdots g_e \cdots)]$ in (32) containing g_e in the argument. The relevant formula is

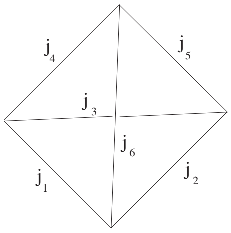
$$P^n_{\mathrm{inv}} := \int \mathrm{d}g \, j_1(g) \otimes j_2(g) \otimes \cdots \otimes j_n(g) = \sum_\iota C^\iota_{j_1 j_2 \cdots j_n} C^{*\iota}_{j_1 j_2 \cdots j_n}, \tag{33}$$

where $P^n_{\rm inv}$ is the projector onto $\mathrm{Inv}[j_1 \otimes j_2 \otimes \cdots \otimes j_n]$. On the RHS we have chosen an orthonormal basis of invariant vectors (intertwiners) to express the projector. Note that the assignment of intertwiners to edges is a consequence of the integration over the connection. This is not a particularity of this example but rather a general property of local spin foams as pointed out in section 3.2. Finally (30) can be written as a sum over spin foam amplitudes

$$\mathcal{Z}(\Delta) = \sum_{\mathcal{C}:\{j\}\rightarrow\{f\}} \sum_{\mathcal{C}:\{\iota\}\rightarrow\{e\}} \prod_{f\in\mathcal{J}_\Delta} \Delta_{j_f} \prod_{v\in\mathcal{J}_\Delta} A_v(\iota_v, j_v), \tag{34}$$

where $A_v(\iota_v, j_v)$ is given by the appropriate trace of the intertwiners ι_v corresponding to the edges bounded by the vertex and j_v are the corresponding representations. This amplitude is given, in general, by an $SU(2)$ $3Nj$-symbol corresponding to the flat evaluation of the spin network defined by the intersection of the corresponding vertex with a 2-sphere. When Δ is a simplicial complex all the edges in $\mathcal{J}_\Delta$ are three-valent and vertices are four-valent (one such vertex is emphasized in figure 3, the intersection with the surrounding S^2 is shown with dotted lines). Consequently, the vertex amplitude is given by the contraction of the corresponding four three-valent intertwiners, i.e., a $6j$-symbol. In that case the partition function takes the familiar Ponzano–Regge [64] form

$$\mathcal{Z}(\Delta) = \sum_{\mathcal{C}:\{j\}\rightarrow\{f\}} \prod_{f\in\mathcal{J}_\Delta} \Delta_{j_f} \prod_{v\in\mathcal{J}_\Delta}$$



, (35)

where the sum over intertwiners disappears since $\dim(\mathrm{Inv}[j_1 \otimes j_2 \otimes j_3]) = 1$ for $SU(2)$ and there is only one term in (33). Ponzano and Regge originally defined the amplitude (35) from the study of the asymptotic properties of the $6j$-symbol.

4.3.1. Discretization independence. A crucial property of the partition function (and transition amplitudes in general) is that it does not depend on the discretization Δ. Given two different cellular decompositions Δ and Δ' (not necessarily simplicial)

$$\tau^{-n_0} Z(\Delta) = \tau^{-n'_0} Z(\Delta'), \tag{36}$$

where n_0 is the number of 0-simplexes in Δ (hence the number of bubbles in $\mathcal{J}_\Delta$), and $\tau = \sum_j (2j+1)$ is clearly divergent which makes discretization independence a formal statement without a suitable regularization.

The sum over spins in (34) is typically divergent, as indicated by the previous equation. Divergences occur due to infinite volume factors corresponding to the topological gauge freedom (24) (see [65]).[13] The factor τ in (36) represents such a volume factor. It can also be interpreted as a $\delta(0)$ coming from the existence of a redundant delta function in (30). One can

[13] For simplicity we concentrate on the Abelian case $G = U(1)$. The analysis can be extended to the non-Abelian case. Writing $g \in U(1)$ as $g = \mathrm{e}^{\mathrm{i}\theta}$ the analogue of the gravity simplicial action is

$$S(\Delta, \{e_f\}, \{\theta_e\}) = \sum_{f\in\mathcal{J}_\Delta} e_f F_f(\{\theta_e\}), \tag{37}$$

where $F_f(\{\theta_e\}) = \sum_{e\in f} \theta_e$. Gauge transformations corresponding to (23) act at the endpoints of edges $e \in \mathcal{J}_\Delta$ by the action of group elements $\{\beta\}$ in the following way:

$$B_f \rightarrow B_f, \qquad \theta_e \rightarrow \theta_e + \beta_s - \beta_t, \tag{38}$$

partially gauge fix this freedom at the level of discretization. This has the effect of eliminating bubbles from the 2-complex.

In the case of simply connected Σ the gauge fixing is complete. One can eliminate bubbles and compute finite transition amplitudes. The result is equivalent to the physical scalar product defined in the canonical picture in terms of the delta measure[14].

In the case of gravity with cosmological constant the state sum generalizes to the Turaev–Viro model [66] defined in terms of $SU_q(2)$ with $q^n = 1$ where the representations are finitely many. Heuristically, the presence of the cosmological constant introduces a physical infrared cut-off. Equation (36) has been proved in this case for the case of simplicial decompositions in [66] (see also [67, 68]). The generalization for arbitrary cellular decomposition was obtained in [69].

4.3.2. Transition amplitudes. Transition amplitudes can be defined along similar lines using a manifold with boundaries. Given Δ, $\mathcal{J}_\Delta$ then defines graphs on the boundaries. Consequently, spin foams induce spin networks on the boundaries. The amplitudes have to be modified concerning the boundaries to have the correct composition property (10). This is achieved by changing the face amplitude from $\left(\Delta_{j_f}\right)$ to $\left(\Delta_{j_\ell}\right)^{1/2}$ on external faces.

The crucial property of this spin foam model is that the amplitudes are independent of the chosen cellular decomposition [67, 69]. This allows for computing transition amplitudes between any spin network states $s = (\gamma, \{j\}, \{\iota\})$ and $s' = (\gamma, \{j'\}, \{\iota'\})$ according to the following rules[15].

- Given $\mathcal{M} = \Sigma \times [0, 1]$ (piecewise linear) and spin network states $s = (\gamma, \{j\}, \{\iota\})$ and $s' = (\gamma, \{j'\}, \{\iota'\})$ on the boundaries—for γ and γ' piecewise linear graphs in Σ—choose any cellular decomposition Δ such that the dual 2-complex $\mathcal{J}_\Delta$ is bordered by the corresponding graphs γ and γ', respectively (existence can be shown easily).
- Compute the transition amplitude between s and s' by summing over all spin foam amplitudes (rescaled as in (36)) for the spin foams $F : s \to s'$ defined on the 2-complex $\mathcal{J}_\Delta$.

4.3.3. The generalized projector. We can compute the transition amplitudes between any element of the kinematical Hilbert space $\mathcal{H}$.[16] Transition amplitudes define the physical scalar product by reproducing the skein relations of the canonical analysis. We can construct the

where the subindex s (respectively t) labels the source vertex (respectively target vertex) according to the orientation of the edge. The gauge invariance of the simplicial action is manifest. The gauge transformation corresponding to (24) acts on vertices of the triangulation Δ and is given by

$$B_f \to B_f + \eta_s - \eta_t, \qquad \theta_e \to \theta_e. \tag{39}$$

According to the discrete analogue of the Stokes theorem

$$\sum_{f \in \mathrm{Bubble}} F_f(\{\theta_e\}) = 0,$$

which implies the invariance of the action under the above transformation. The divergence of the corresponding spin foam amplitudes is due to this last freedom. Alternatively, one can understand it from the fact that Stokes theorem implies a redundant delta function in (30) per bubble in $\mathcal{J}_\Delta$ [65].

[14] If $\mathcal{M} = S^2 \times [0, 1]$ one can construct a cellular decomposition interpolating any two graphs on the boundaries without having internal bubbles and hence no divergences.

[15] Here we are ignoring various technical issues in order to emphasize the relevant ideas. The most delicate is that of the divergences due to gauge factors mentioned above. For a more rigorous treatment, see [70].

[16] The sense in which this is achieved should be apparent from our previous definition of transition amplitudes. For a rigorous statement, see [70].

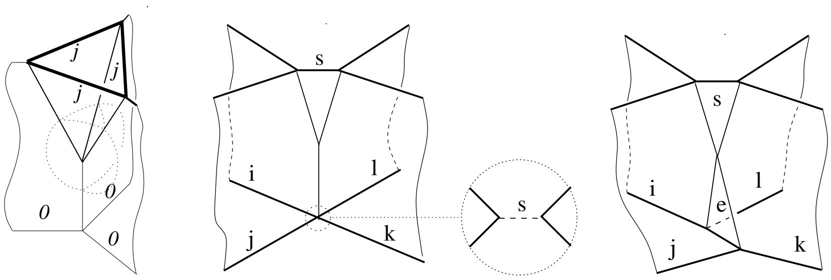


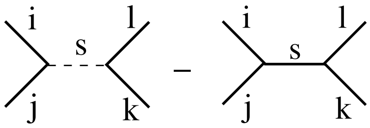
Figure 5. Elementary spin foams used to prove skein relations.

physical Hilbert space by considering equivalence classes under states with zero transition amplitude with all the elements of $\mathcal{H}$, i.e., null states.

Here we explicitly construct a few examples of null states. For any contractible Wilson loop in the j representation the state

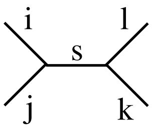
$$\psi = (2j+1)s - \bigcirc_{j} \otimes s \underset{\text{phys}}{=} 0, \qquad (40)$$

for any spin network state s, has vanishing transition amplitude with any element of $\mathcal{H}$. This can be easily checked by using the rules stated above and the portion of spin foam illustrated in figure 5 to show that the two terms in the previous equation have the same transition amplitude (with opposite sign) for any spin network state in $\mathcal{H}$. Using the second elementary spin foam in figure 5, one can similarly show that

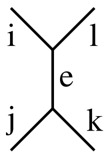


$$\underset{\text{phys}}{=} 0, \qquad (41)$$

or the re-coupling identity using the elementary spin foam on the right of figure 5



$$- \sum_{e} \sqrt{2s+1}\sqrt{2e+1} \begin{Bmatrix} i & j & s \\ k & l & e \end{Bmatrix}$$



$$\underset{\text{phys}}{=} 0, \qquad (42)$$

where the quantity in brackets represents an $SU(2)6j$-symbol. All skein relations can be found in this way. The transition amplitudes imply the skein relations that define the physical Hilbert space! The spin foam quantization is equivalent to the canonical one.

4.3.4. The continuum limit. Recently Zapata [70] formalized the idea of a continuum spin foam description of three-dimensional gravity using projective techniques inspired by those utilized in the canonical picture [24]. The heuristic idea is that due to the discretization invariance one can define the model in an 'infinitely' refined cellular decomposition that contains any possible spin network state on the boundary (this intuition is implicit in our rules for computing transition amplitudes mentioned above). Zapata concentrates on the case with non-vanishing cosmological constant and constructs the continuum extension of the Turaev–Viro model.

4.4. Conclusion

We have illustrated the general notion of the spin foam quantization in the simple case of three-dimensional Riemannian gravity (for the generalization to the Lorentzian case, see [60]). The main goal of the approach is to provide a definition of the physical Hilbert space. The example of this section sets the guiding principles of what one would like to realize in four dimensions. However, as should be expected, there are various new issues that make the task by far more involved.

5. Spin foams for four-dimensional quantum gravity

In this section, we briefly describe the various spin foam models for quantum gravity in the literature.

5.1. The Reisenberger model

According to Plebanski [58] the action of self-dual Riemannian gravity can be written as a constrained $SU(2)$ BF theory

$$S(B,A) = \int \mathrm{Tr}[B \wedge F(A)] - \psi_{ij}\left[B^i \wedge B^j - \frac{1}{3}\delta^{ij} B^k \wedge B_k\right], \tag{43}$$

where variations with respect to the symmetric (Lagrange multiplier) tensor ψ_{ij} imposes the constraints

$$\Omega^{ij} = B^i \wedge B^j - \tfrac{1}{3}\delta^{ij} B^k \wedge B_k = 0. \tag{44}$$

When B is non-degenerate the constraints are satisfied if and only if $B^i = \pm\big(e^0 \wedge e^i + \frac{1}{2}\epsilon^i_{\ jk} e^j \wedge e^k\big)$ which reduces the previous action to that of self-dual general relativity. Reisenberger studied the simplicial discretization of this action in [56] as a preliminary step towards the definition of the corresponding spin foam model. The consistency of the simplicial action is argued by showing that the simplicial theory converges to the continuum formulation when the triangulation is refined: both the action and its variations (equations of motion) converge to those of the continuum theory.

In [57], Reisenberger constructs a spin foam model for this simplicial theory by imposing the constraints Ω^{ij} directly on the $SU(2)$ BF amplitudes. The spin foam path integral for BF theory is obtained as in section 4. The constraints are imposed by promoting the B^i to operators χ^i (combinations of left-/right-invariant vector fields) acting on the discrete connection[17]. The model is defined as

$$Z_{GR} = \underbrace{\int \prod_{e\in\mathcal{J}_\Delta} \mathrm{d}g_e}_{\int \mathcal{D}[A]} \; \overbrace{\delta(\hat{\Omega}^{ij})}^{\int \mathcal{D}[\psi]\, \mathrm{e}^{\mathrm{i}\psi_{ij}\Omega^{ij}}} \; \underbrace{\sum_{\mathcal{C}:\{j\}\to\{f\}} \prod_{f\in\mathcal{J}_\Delta} \Delta_{j_f} \mathrm{Tr}\left[j_f\left(g_e^1 \cdots g_e^N\right)\right]}_{\int \mathcal{D}[B]\, \mathrm{e}^{\mathrm{i}\int \mathrm{Tr}[B\wedge F(A)]}}, \tag{46}$$

where $\hat{\Omega} = \mathcal{J}^i \wedge \mathcal{J}^j - \frac{1}{3}\delta^{ij}\mathcal{J}^k \wedge \mathcal{J}_k$ and we have indicated the correspondence of the different terms with the continuum formulation. The preceding equation is rather formal;

[17] Note that (for example) the right-invariant vector field $\mathcal{J}^i(U) = \sigma^{iA}_B U^B_C\, \partial/\partial U^A_C$ has a well-defined action at the level of equation (32) and acts as a B operator at the level of (29) since

$$-\mathrm{i}\chi^i(U)[\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}]|_{U\sim 1} = \mathrm{Tr}[\sigma^i U B]\, \mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}|_{U\sim 1} \sim B^i\, \mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}, \tag{45}$$

where σ^i are Pauli matrices.

for the rigorous implementation, see [57]. Reisenberger uses locality so that constraints are implemented on a single 4-simplex amplitude. There is, however, a difficulty with this procedure: the algebra of operators $\hat{\Omega}^{ij}$ does not close so that imposing the constraints sharply becomes too a strong condition on the BF configurations[18]. In order to avoid this, Reisenberger defines a one-parameter family of models by inserting the operator

$$\mathrm{e}^{-\frac{1}{2z^2}\hat{\Omega}^2} \tag{47}$$

instead of the delta function above. In the limit $z \to \infty$ the constraints are sharply imposed. This introduces an extra parameter to the model. The properties of the kernel of $\hat{\Omega}$ have not been studied in detail.

5.2. The Freidel–Krasnov prescription

In [71], Freidel and Krasnov define a general framework to construct spin foam models corresponding to theories whose action has the general form

$$S(B, A) = \int \mathrm{Tr}[B \wedge F(A)] + \Phi(B), \tag{48}$$

where the first term is the BF action while $\Phi(B)$ is a certain polynomial function of the B field. The formulation is constructed for compact internal groups. The definition is based on the formal equation

$$\int \mathcal{D}[B]\mathcal{D}[A]\, \mathrm{e}^{\mathrm{i}\int \mathrm{Tr}[B\wedge F(A)]+\Phi(B)} := \mathrm{e}^{\mathrm{i}\int \Phi(\frac{\delta}{\delta J})} Z[J]|_{J=0}, \tag{49}$$

where the *generating functional* $Z[J]$ is defined as

$$Z[J] := \int \mathcal{D}[B]\mathcal{D}[A]\, \mathrm{e}^{\mathrm{i}\int \mathrm{Tr}[B\wedge F(A)]+\mathrm{Tr}[B\wedge J]}, \tag{50}$$

where J is an algebra-valued 2-form field. They provide a rigorous definition of the generating functional by introducing a discretization of $\mathcal{M}$ in the same spirit as the other spin foam models discussed here. Their formulation can be used to describe various theories of interest such as BF theories with cosmological terms, Yang–Mills theories (in two dimensions) and Riemannian self-dual gravity. In the case of self-dual gravity B and A are valued in $su(2)$, while

$$\Phi(B) = \int \psi_{ij} \left[B^i \wedge B^j - \frac{1}{3}\delta^{ij} B^k \wedge B_k \right], \tag{51}$$

according to equation (43). The model obtained in this way is very similar to Reisenberger's one. There are, however, some technical differences. One of the more obvious ones is that the non-commutative invariant vector fields $\mathcal{J}^i$ representing B^i are replaced here by the commutative functional derivatives $\delta/\delta J^i$. The explicit properties of these models have not been studied further.

5.3. The Iwasaki model

Iwasaki defines a spin foam model of self-dual Riemannian gravity[19] by a direct lattice discretization of the continuous Ashtekar formulation of general relativity. This model constitutes an example of the general prescription of section 3.2. The action is

$$S(e, A) = \int \mathrm{d}x^4 \epsilon^{\mu\nu\lambda\sigma} \left[2e^0_{[\mu} e_{\nu]i} + \epsilon^0_{ijk} e^j_\mu e^k_\nu\right] \left[2\partial_{[\lambda} A^i_{\sigma]} + \epsilon^{0i}{}_{lm} A^l_\lambda A^m_\sigma\right], \tag{52}$$

[18] This difficulty also arises in the Barrett–Crane model as we shall see in section 7.

[19] Iwasaki defines another model involving multiple cellular complexes to provide a simpler representation of wedge products in the continuum action. A more detailed presentation of this model would require the introduction of various technicalities at this stage, so we refer the reader to [72].

where A_a^i is an $SU(2)$ connection. The fundamental observation of [73] is that one can write the discrete action in a very compact form if we encode part of the degrees of freedom of the tetrad in an $SU(2)$ group element. More precisely, if we take $g_{\mu\mu} = e_\mu^i e_\mu^j \delta_{ij} = 1$ we can define $\mathbf{e}_\mu := e_\mu^0 + \mathrm{i}\sigma_i e_\mu^i \in SU(2)$ where σ_i are the Pauli matrices. In this parametrization of the 'angular' components of the tetrad and using a hypercubic lattice the discrete action becomes

$$S_\Delta = -\beta \sum_{v\in\Delta} \epsilon^{\mu\nu\lambda\sigma} r_\mu r_\nu \operatorname{Tr}\left[\mathbf{e}_\mu^\dagger \mathbf{e}_\nu U_{\lambda\sigma}\right], \tag{53}$$

where $r_\mu := (\beta^{1/2}\ell_p)^{-1}\epsilon\sqrt{g_{\mu\mu}}$, $U_{\mu\nu}$ is the holonomy around the $\mu\nu$-plaquette, ϵ the lattice constant and β is a cut-off for r_μ used as a regulator ($r_\mu \leqslant \beta^{1/2}\ell_p\epsilon^{-1}$). The lattice path integral is defined by using the Haar measure both for the connection and the 'spherical' part of the tetrad $\mathbf{e}$ and the radial part $\mathrm{d}r_\mu := \mathrm{d}r_\mu r_\mu^3$. The key formula to obtain an expression involving spin foams is

$$\mathrm{e}^{\mathrm{i}x\operatorname{Tr}[U]} = \sum_j (2j+1)\frac{J_{2j+1}(2x)}{x}\chi_j(U). \tag{54}$$

Iwasaki writes down an expression for the spin foam amplitudes in which the integration over the connection and the $\mathbf{e}$ can be computed explicitly. Unfortunately, the integration over the radial variables r involves products of Bessel functions and its behaviour is not analysed in detail. In three dimensions the radial integration can be done and the corresponding amplitudes coincide with the results of section 4.3.

5.4. The Barrett–Crane model

The appealing feature of the previous models is the clear connection to loop quantum gravity, since they are defined directly using the self-dual formulation of gravity (boundary states are $SU(2)$ spin networks). The drawback is the lack of closed simple expressions for the amplitudes which complicates their analysis. There is, however, a simple model that can be obtained as a systematic quantization of the simplicial $SO(4)$ Plebanski's action. This model was introduced by Barrett and Crane in [9] and further motivated by Baez in [14]. The basic idea behind the definition was that of the *quantum tetrahedron* introduced by Barbieri in [74] and generalized to 4D in [15]. The beauty of the model resides in its remarkable simplicity. This has stimulated a great deal of explorations and produced many interesting results. We will review most of these in section 7.

5.5. Markopoulou–Smolin causal spin networks

Using the kinematical setting of LQG with the assumption of the existence of a micro-local (in the sense of Planck scale) causal structure Markopoulou and Smolin define a general class of (causal) spin foam models for gravity [75, 76] (see also [77]). The elementary transition amplitude $A_{s_I\to s_{I+1}}$ from an initial spin network s_I to another spin network s_{I+1} is defined by a set of simple combinatorial rules based on a definition of causal propagation of the information at nodes. The rules and amplitudes have to satisfy certain causal restrictions (motivated by the standard concepts in classical Lorentzian physics). These rules generate surface-like excitations of the same kind we encounter in the more standard spin foam model but endow the foam with a notion of causality. Spin foams $\mathcal{F}_{s_i\to s_f}^N$ are labelled by the number of times these elementary transitions take place. Transition amplitudes are defined as

$$\langle s_i, s_f\rangle = \sum_N \prod_{I=0}^{N-1} A_{s_I\to s_{I+1}}. \tag{55}$$

The models are not related to any continuum action. The only guiding principles are the restrictions imposed by causality, simplicity and the requirement of the existence of a non-trivial critical behaviour that would reproduce general relativity at large scales. Some indirect evidence of a possible non-trivial continuum limit has been obtained in some versions of the model in 1 + 1 dimensions.

5.6. Gambini–Pullin model

In [78] Gambini and Pullin introduced a very simple model obtained by modification of the BF theory skein relations. As we argued in section 4 skein relations defining the physical Hilbert space of BF theory are implied by the spin foam transition amplitudes. These relations reduce the large kinematical Hilbert space of BF theory (analogous to that of quantum gravity) to a physical Hilbert space corresponding to the quantization of a finite number of degrees of freedom. Gambini and Pullin define a model by modifying these amplitudes so that some of the skein relations are now forbidden. This simple modification frees local excitations of a field theory. A remarkable feature is that the corresponding physical states are (in a certain sense) solutions to various regularizations of the scalar constraint for (Riemannian) LQG. The fact that physical states of BF theory solve the scalar constraint is well known [79], since roughly $F(A) = 0$ implies $EEF(A) = 0$. The situation here is of a similar nature, and, as the authors argue, one should interpret this result as an indication that some 'degenerate' sector of quantum gravity might be represented by this model. The definition of this spin foam model is not explicit since the theory is directly defined by the physical skein relations.

5.7. Capovilla–Dell–Jacobson theory on the lattice

The main technical difficulty that we gain in going from three-dimensional general relativity to the four-dimensional one is that the integration over the e becomes intricate. In the Capovilla–Dell–Jacobson [80, 81] formulation of general relativity this 'integration' is partially performed at the continuum level. The action is

$$S(\eta, A) = \int \eta\, \mathrm{Tr}[\epsilon \cdot F(A) \wedge F(A) \epsilon \cdot F(A) \wedge F(A)], \tag{56}$$

where $\epsilon \cdot F \wedge F := \epsilon^{abcd} F_{ab} F_{cd}$. Integration over η can be formally performed in the path integral and we obtain

$$Z = \int \prod_x \delta(\mathrm{Tr}[\epsilon \cdot F(A) \wedge F(A) \epsilon \cdot F(A) \wedge F(A)]). \tag{57}$$

This last expression is of the form (17) and can be easily discretized along the lines of section 3.2. The final expression (after integrating over the lattice connection) involves a sum over spin configurations with no implicit integrations. One serious problem of this formulation is that it corresponds to a sector of gravity where the Weyl tensor satisfies certain algebraic requirements. In particular, flat geometries are not contained in this sector.

6. Some conceptual issues

6.1. Anomalies and gauge fixing

As we mentioned before and illustrated with the example of section 4, the spin foam path integral is meant to provide a definition of the physical Hilbert space. Spin foam transition amplitudes are not interpreted as defining propagation in time but rather as defining the physical scalar product. This interpretation of spin foam models is the only one consistent with general covariance. However, in the path integral formulation, this property relies on the gauge invariance of the path integral measure. If the measure meets this property we say it is *anomaly free*. It is well known that in addition to the invariance of the measure, one must provide appropriate gauge-fixing conditions for the amplitudes to be well defined. In this section, we analyse these issues in the context of the spin foam approach.

Since we are interested in gravity in the first-order formalism, in addition to diffeomorphism invariance one has to deal with the gauge transformations in the internal space. Let us first describe the situation for the latter. If this gauge group is compact then anomaly-free measures are defined using appropriate variables and invariant measures. In this case gauge fixing is not necessary for the amplitudes to be well defined. Examples where this happens are: the models of Riemannian gravity considered in this topical review (the internal gauge group being $SO(4)$ or $SU(2)$), and standard lattice gauge theory. In these cases, one represents the connection in terms of group elements (holonomies) and uses the (normalized) Haar measure in the integration. In the Lorentzian sector (internal gauge group $SL(2,\mathbb{C})$) the internal gauge orbits have infinite volume and the lattice path integral would diverge without an appropriate gauge-fixing condition. These conditions generally exist in spin foam models and we will study an example in section 9 (for a general treatment, see [82]).

The remaining gauge freedom is diffeomorphism invariance. It is generally assumed that spin foam encodes diff-invariant information about 4-geometries. Thus, no gauge fixing would be necessary as one would be already summing over physical configurations. To illustrate this perspective we concentrate on the case of a model defined on a fixed discretization Δ as described in section 3.2.

Let us start by considering the spin network states, at $\partial\Delta$: boundary of Δ, for which we want to define the transition amplitudes. According to what we have learned from the canonical approach, 3-diffeomorphism invariance is implemented by considering (diffeomorphism) equivalence classes of spin network states. In the context of spin foams, the underlying discretization Δ restricts the graphs on the boundary to be contained on the dual 1-skeleton of the boundary complex $\partial\Delta$. These states are regarded as representative elements of the corresponding 3-diffeomorphism equivalence class. The discretization can be interpreted, in this way, as a gauge fixing of 3-diffeomorphisms on the boundary. This gauge fixing is partial in the sense that, generically, there will remain a discrete symmetry remnant given by the discrete symmetries of the spin network. This remaining symmetry has to be factored out when computing transition amplitudes (in fact this also plays a role in the definition of the kinematical Hilbert space of section 2).

The standard view point (consistent with LQG and quantum geometry) is that this should naturally generalize to 4-diffeomorphisms for spin foams. The underlying 2-complex $\mathcal{J}_\Delta$ on which spin foams are defined represents a partial gauge fixing for the configurations (spin foams) entering the path integral. The remaining symmetry, to be factored out in the computation of transition amplitudes, corresponds simply to the finite group of discrete symmetries of the corresponding spin foams[20]. This factorization is well defined since the number of equivalent spin foams can be characterized in a fully combinatorial manner, and is finite for any spin foam defined on a finite discretization. In addition, a spin foam model is anomaly free if the amplitudes are invariant under this discrete symmetry. This requirement is met by all the spin foam models we considered in this topical review.

We illustrate the intuitive idea of the previous paragraph in figure 6. In diagram (a) a continuum configuration is represented by a discrete (spin foam) configuration on the lattice. In (b) two configurations are shown. In the background-dependent context (e.g., lattice gauge

[20] Baez [14] points out this equivalence relation between spin foams as a necessary condition for the definition of the *category of spin foams*.

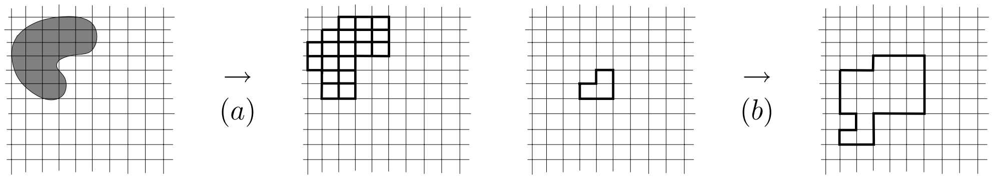


Figure 6. Diffeomorphisms are replaced by discrete symmetries in spin foams. In this simplified picture there are two possible colourings of 0 or 1, the latter represented by the darkening of the corresponding lattice element. The two configurations in (*b*) are regarded as physically equivalent in the background-independent context.

theory) these two configurations would be physically inequivalent as the lattice carries metric information (length of the edges). In the context of spin foam models there is no geometric information encoded in the discretization and in an anomaly-free spin foam model the two configurations should be regarded as equivalent.

The discretization of the manifold Δ is seen as a regulator introduced to define the spin foam model. Even when the regulator (or the discretization dependence) eventually has to be removed (see the following subsection), the theory remains discrete at the fundamental level. The smooth manifold diffeomorphism-invariant description is regarded in this context as a derived concept in the (low-energy) continuum limit. Fundamental excitations are intrinsically discrete. From this viewpoint, the precise meaning of the gauge symmetries of such a theory would have to be formulated directly at the discrete level. We have seen that this can be achieved in the case of three-dimensional gravity (recall section 4.3.1). From this perspective, the discrete symmetries of coloured 2-complexes (spin foams) represent the fundamental 'gauge' freedom. This symmetry would manifest itself as diffeomorphisms only in the continuum limit.

The viewpoint stated above is consistent with recent results obtained by Gambini and Pullin [83, 84]. They study the canonical formulation of theories defined on a lattice from the onset. This provides a way to analyse the meaning of gauge symmetries directly in the style of Dirac. Their results indicate that diff-invariance is indeed broken by the discretization in the sense that there is no infinitesimal generator of diffeomorphism. This is consistent with our covariant picture of discrete symmetries above. In their formulation the canonical equations of motion fix the value of what were Lagrange multipliers in the continuum (e.g., lapse and shift). This is interpreted as a breaking of diffeomorphism invariance; however, the solutions of the multiplier equations are highly non-unique. The ambiguity in selecting a particular solution corresponds to the remnant diffeomorphism invariance of the discrete theory.

However, the notion of anomaly freeness stated at the beginning of this section should be strengthened. In fact, according to our tentative definition, an anomaly-free measure can be multiplied by any gauge-invariant function and yield a new anomaly-free measure. This kind of ambiguity is not what we want; however, it is, in fact, present in most of the spin foam models defined so far. In standard QFT theory, the formal (phase space) path integral measure in the continuum has a unique meaning (up to a constant normalization) emerging from the canonical formulation. Provided an appropriate gauge fixing, the corresponding Dirac bracket determines the formal measure on the gauge-fixed constraint surface. As we will see later, there is a certain degree of ambiguity in the definition of various spin foam models in four dimensions. This ambiguity concerns the evaluation of lower-dimensional simplexes and is directly related to the definition of measure in the spin foam sum. One would expect that a strengthened definition of anomaly freeness should resolve these ambiguities. This possibility is studied in [85].

Finally, let us briefly recall the situation in three-dimensional gravity. In three dimensions the discrete action is invariant under transformations that are in correspondence with the continuum gauge freedoms (23) and (24). As we mentioned in section 4.3.1, spin foam amplitudes diverge due to (24). This means that spin foams do not fix (24) up to a 'finite volume' discrete symmetry, which seems to be in conflict with the argument stated above as diffeomorphism can be obtained combining the transformations (23) and (24). However, the topological gauge symmetry (24) involves more than just diffeomorphisms. Indeed, only on shell can one express diffeomorphisms as a combination of (23) and (24). This representation of diffeomorphisms turns out to be field dependent (recall footnote 11). At the canonical level we have seen that the topological symmetry (24) acts in a way that is totally different from diffeomorphisms. In particular, spin network states defined on graphs which differ in the number of edges and vertices can be physically equivalent (recall section 4.3.3). Thus three-dimensional gravity is a degenerate example and extrapolation of its gauge properties to four dimensions seems misleading.

In section 7.1, we will study the spin foam quantization of BF theory—a topological theory which corresponds to a generalization of three-dimensional gravity to four dimensions. The spin foam amplitudes are also divergent in this case due to the analogue of the gauge symmetry (24). Some of the spin foam models in section 5 are defined from BF theory by implementing constraints that reduce the topological theory to general relativity. The implementation of the constraints breaks the topological symmetry (24) and the resulting model is no longer topological invariant. At the continuum level it is clear that the remnant gauge symmetry is diffeomorphism invariance. Whether the resulting spin foam model for gravity remnant is larger than the discrete spin foam symmetries advocated above has not been studied in detail.

The action of diffeomorphism is far from understood at a rigorous level in the context of spin foam models. Here we have presented an account of some of the ideas under consideration, and we tried to point out the relevance of an issue that certainly deserves detailed investigation.

6.2. Discretization dependence

The spin foam models we have introduced so far are defined on a fixed cellular decomposition of $\mathcal{M}$. This is to be interpreted as an intermediate step towards the definition of the theory. The discretization reduces the infinite-dimensional functional integral to a multiple integration over a finite number of variables. This cut-off is reflected by the fact that only a restrictive set of spin foams (spin network histories) is allowed in the path integral: those that can be obtained by all possible colourings of the underlying 2-complex. In addition, it restricts the number of possible 3-geometry states (spin network states) on the boundary by fixing a finite underlying boundary graph. This represents a truncation in the allowed fluctuations and the set of states of the theory that can be interpreted as a regulator. However, the nature of this regulator is fundamentally different from the standard concept in the background-independent framework: since geometry is encoded in the colouring (that can take any spin values) the configurations involve fluctuations all the way to the Planck scale[21]. This scenario is different in lattice gauge theories where the lattice introduces an effective UV cut-off given by the lattice spacing. Transition amplitudes are, however, discretization dependent now. A consistent definition of the path integral using spin foams should include a prescription for eliminating this discretization dependence.

A special case is that of topological theories such as gravity in three dimensions. In this case, one can define the sum over spin foams with the aid of a fixed cellular decomposition

[21] Changing the label of a face from j to $j+1$ amounts to changing an area eigenvalue by an amount of the order of Planck length squared according to (7).

Δ of the manifold. Since the theory has no local excitations (no gravitons), the result is independent of the chosen cellular decomposition. A single discretization suffices to capture the degrees of freedom of the topological theory.

In lattice gauge theory the solution to the problem is implemented through the so-called continuum limit. In this case the existence of a background geometry is crucial, since it allows one to define the limit when the lattice constant (length of links) goes to zero. In addition, the possibility of working in the Euclidean regime allows the implementation of statistical mechanical methods.

None of these structures is available in the background-independent context. The lattice (triangulation) contains only topological information and there is no geometrical meaning associated with its components. As we mentioned above this has the novel consequence that the truncation cannot be regarded as a UV cut-off as in the background-dependent context. This, in turn, represents a conceptual obstacle to the implementation of standard techniques. Moreover, no Euclidean formulation seems meaningful in a background-independent scenario. New means to eliminate the truncation introduced by the lattice have to be developed.

This is a major issue where concrete results have not been obtained so far beyond the topological case. Here we explain the two main approaches to recover general covariance corresponding to the realization of the notion of 'summing over discretizations' of [54].

- *Refinement of the discretization.* According to this idea topology is fixed by the simplicial decomposition. The truncation in the number of degrees of freedom should be removed by considering the triangulations of an increasing number of simplexes for that fixed topology. The flow in the space of possible triangulations is controlled by the Pachner moves. The formal idea is to take a limit in which the number of 4-simplexes goes to infinity together with the number of tetrahedra on the boundary. Given a 2-complex $\mathcal{J}_2$ which is a refinement of a 2-complex $\mathcal{J}_1$ then the set of all possible spin foams defined on $\mathcal{J}_1$ is naturally contained in those defined on $\mathcal{J}_2$ (taking into account the equivalence relations for spin foams mentioned in the previous section). The refinement process should also enlarge the space of possible 3-geometry states (spin networks) on the boundary recovering the full kinematical sector in the limit of infinite refinements. An example where this procedure is well defined is Zapata's treatment of the Turaev–Viro model [70]. The key point in this case is that amplitudes are independent of the discretization (due to the topological character of the theory) so that the refinement limit is trivial. In the general case there is a great deal of ambiguity involved in the definition of refinement[22]. The hope is that the nature of the transition amplitudes would be such that these ambiguities will not affect the final result. The Turaev–Viro model is an example where this prescription works.

  If the refinement limit is well defined one would expect that working with a 'sufficiently refined' but fixed discretization would serve as an approximation that could be used to extract physical information with some quantifiable precision.

[22] It is not difficult to define the refinement in the case of a hypercubic lattice. In the case of a simplicial complex a tentative definition can be attempted using Pachner moves. To illustrate we can concentrate on the simple two-dimensional case. Given an initial triangulation of a surface with boundary Δ_1 define Δ_1' by implementing a 1–3 Pachner move to each triangle in Δ_1. Δ_1' is a homogeneous refinement of Δ_1; however, the boundary triangulation remains unchanged so that they will support the same space of boundary data or spin networks. As mentioned above, in the refinement process one also wants to refine the boundary triangulation so that the corresponding dual graph will get refined and the space of possible boundary data will become larger (it should involve all of Cyl in the limit). In order to achieve this we define Δ_2 by erasing from Δ_1' triangles sharing a 1-simplex with the boundary. This amounts to carrying out a 1–2 Pachner move on the boundary. This completes the refinement $\Delta_1 \to \Delta_2$. This refinement procedure seems, however, not fully satisfactory as the refinement keeps the memory of the initial triangulation. This can be easily visualized in the dual 2-complex. An improvement of this prescription could involve the random implementation of 2–2 Pachner moves between refinements.

- *Spin foams as Feynman diagrams.* This idea has been motivated by the generalized matrix models of Boulatov and Ooguri [86, 87]. The fundamental observation is that spin foams admit a dual formulation in terms of a field theory over a group manifold [11, 16, 17]. The duality holds in the sense that spin foam amplitudes correspond to Feynman diagram amplitudes of the GFT. The perturbative Feynman expansion of the GFT (expansion in a fiducial coupling constant λ) provides a definition of *sum over* discretizations which is fully combinatorial and hence independent of any manifold structure[23]. The latter is the most appealing feature of this approach.

  However, the convergence issues clearly become more involved. The perturbative series are generically divergent. This is not necessarily a definite obstruction as divergent series can often be given an asymptotic meaning and provide physical information. Moreover, there are standard techniques that can allow us to 're-sum' a divergent series in order to obtain non-perturbative information. Recently, Freidel and Louapre [89] have shown that this is indeed possible for certain GFTs in three dimensions. Other possibilities have been proposed in [16].

  It is not clear how the notion of diffeomorphism would be addressed in this framework. Diffeomorphism-equivalent configurations (in the discrete sense described above) appear at all orders in the perturbation series[24]. From this perspective (and leaving aside the issue of convergence) the sum of different order amplitudes corresponding to equivalent spin foams should be interpreted as the definition of the physical amplitude. The discussion of the previous section does not apply in the GFT formulation, i.e., there is no need for gauge fixing.

  The GFT formulation resolves by definition the two fundamental conceptual problems of the spin foam approach: diffeomorphism gauge symmetry and discretization dependence. The difficulties are shifted to the question of the physical role of λ and the convergence of the corresponding perturbative series. The GFT formulation has also been very useful in the definition of a simple normalization of the Barret–Crane model and has simplified its generalizations to the Lorentzian sector. We will study this formulation in detail in section 8.

6.3. *Physical scalar product revisited*

In this subsection, we study the properties of the generalized projection operator P introduced before. The generalized projection operator is a linear map from the kinematical Hilbert space into the physical Hilbert space (states annihilated by the scalar constraint). These states are not contained in the kinematical Hilbert space but are rather elements of the dual space Cyl^*. For this reason the operator P^2 is not well defined. We have briefly mentioned the construction of the generalized projection operator in the particular context of Rovelli's model of section 3.1. After the above discussion of gauge and discretization dependence we want to revisit the construction of the P operator. We consider the case in which discretization independence is obtained by a refinement procedure.

The number of 4-simplexes N in the triangulation plays the role of cut-off. The matrix elements of the P operator are expected to be defined by the refinement limit of fixed discretization transition amplitudes ($\langle sP_N, s'\rangle$), namely

$$\langle s, s'\rangle_{\text{phys}} = \langle sP, s'\rangle = \lim_{N\to\infty} \langle sP_N, s'\rangle. \tag{58}$$

[23] This is more than a 'sum over topologies' as many of the 2-complexes appearing in the perturbative expansion cannot be associated with any manifold [88].

[24] The GFT formulation is clearly non-trivial already in the case of topological theories. There have been attempts to make sense of the GFT formulation dual to BF theories in lower dimensions [90].

If we have a well-defined operator P (formally corresponding to $\delta[\hat{\mathcal{S}}]$) the reconstruction of the physical Hilbert space goes along the lines of the GNS construction [91] where the algebra is given by Cyl and the state $\omega(\Psi)$ for $\Psi \in Cyl$ is defined by P as

$$\omega(\Psi) = \frac{\langle 0P, \Psi\rangle}{\langle 0P, 0\rangle}.$$

In the spin foam language this corresponds to the transition amplitude from the state Ψ to the 'vacuum'. The representation of this algebra implied by the GNS construction has been used in [92] to define observables in the theory.

In section 3.1, we mentioned another way to obtain observables according to the definition

$$\langle s O_{\rm phys}, s'\rangle = \langle s P O_{\rm kin} P, s'\rangle, \tag{59}$$

where $O_{\rm kin}$ is a kinematical observable. This construction is expected to be well defined only for some suitable kinematical operators $O_{\rm kin}$ (e.g., the previous equation is clearly divergent for the kinematical identity). There are kinematical operators which can be used as regulators[25]. They correspond to Rovelli's 'sufficiently localized in time' operators [54]. In [93] it is argued that in posing a physical question one is led to conditioning the path integral and in this way improving the convergence properties. Conditioning is represented here by the regulator operators (see the previous footnote).

6.4. Contact with the low-energy world

A basic test of any theory of quantum gravity is the existence of a well-defined classical limit corresponding to general relativity. In the case of the spin foam approach this should be

[25] Consider the following simple example. Take the torus $S^1 \times S^1$ represented by $[0, 2\pi] \times [0, 2\pi] \in \mathrm{R}^2$ with periodic boundary conditions. An orthonormal basis of states is given by the ('spin network') wavefunctions $\langle xy, nm\rangle = \frac{1}{2\pi}\, \mathrm{e}^{\mathrm{i}nx}\, \mathrm{e}^{\mathrm{i}my}$ for $n, m \in \mathbb{Z}$. Take the scalar constraint to be given by $\hat{X}$ so that the generalized projection P simply becomes

$$\langle nm, \hat{P}| = \sum_\alpha \langle \alpha m|.$$

The RHS is no longer in $\mathcal{H}_{kin}$ since it is not normalizable: it is meaningful as a distributional state. The physical scalar product is defined as in (58), namely

$$\langle nm, n'm'\rangle_{\rm phys} = \langle nmP, n'm'\rangle = \sum_\alpha \langle \alpha m, n'm'\rangle = \delta_{mm'}.$$

Therefore, we have that $|nm\rangle \underset{\rm phys}{=} |km\rangle \forall k$. We can define physical observables by means of the formula

$$\langle m, \hat{O}_{ph}|m'\rangle = \langle nm, \hat{P}\hat{O}\hat{P}|n'm'\rangle,$$

for suitable kinematic observables O. If, for example, we take $O = 1$ the previous equation would involve $\hat{P}^2$ which is not defined. However, take the following family of kinematical observables

$$|O_k, nm\rangle = \delta_{n,k}|km\rangle,$$

then $\langle m, \hat{O}_{\rm ph}|m'\rangle = \delta_{m,m'}$, i.e., $\hat{O}_{\rm ph}$ is just the physical identity operator for any value of k. The whole family of kinematic operators defines a single physical operator. The key to the convergence of the previous definition is in the fact that O_k is a projection operator into an eigenstate of the operator $\hat{P}_x$ canonically conjugate to the constraint $\hat{X}$. In particular, any operator that decays sufficiently fast going to $\pm\infty$ in the spectrum of $\hat{P}_x$ will do. This illustrates what kind of operators will be suitable for the above definition to work: they have to be sufficiently localized along the gauge orbit, hence 'sufficiently localized in time' in the case of the scalar constraint. For example, one may also use the Gaussian $\hat{O}_k|nm\rangle = \mathrm{e}^{\frac{-(k-n)^2}{\epsilon}}|nm\rangle$ which also projects down to a multiple of the physical identity. Moreover, the availability of such a kinematical operator can serve as a regulator of non-suitable operators, for instance

$$\langle m, \hat{O}_{\rm ph}|m'\rangle = \langle nm, \hat{P}\hat{O}_k\hat{O}\hat{O}_k\hat{P}|n'm'\rangle,$$

will converge even for $\hat{O} = 1$. The operator $\hat{O}_k$ can be regarded as a quantum gauge fixing: it selects a point along the gauge orbit generated by the constraint.

accomplished simultaneously with a limiting procedure that bridges the fundamental discrete theory with the smooth description of classical physics. This operation is sometimes referred to as the *continuum limit* but it should not be confused with the issues analysed in the previous subsection.

There is debate on how one would actually set up the definition of the 'low-energy limit' in the background-independent context. The general strategy—in fact motivated by our experience in background physics—is to set up a renormalization scheme (in the style of Wilson) where microscopic degrees of freedom are summed out to obtain a coarse-grained effective description. The previous strategy is to be regarded as heuristic until a clear-cut definition is found which encompasses all the issues discussed in this section. It should be emphasized that even when various spin foam models for quantum gravity have been defined, none of them has been shown to reproduce gravity at 'low energies'. This is a major open question that deserves all our efforts.

Interesting ideas on how to define the renormalization programme for spin foams have been proposed by Markopoulou [94, 95]. They make use of novel mathematical techniques shown to be useful in dealing with Feynman diagrams in standard QFT [96, 97]. In this approach Pachner moves on the underlying 2-complex are regarded as radiative corrections in standard QFT[26]. In order to understand the 'scaling' properties of the theory one has to study the behaviour of amplitudes under the action of these elementary moves. Amplitudes for these elementary moves in the Barrett–Crane model have been computed in [12]. A full implementation of Markopoulou's proposal to a concrete model remains to be studied.

Recently Oeckl [98] has studied the issue of renormalization in the context of spin foam models containing a coupling parameter. These models include generalized covariant gauge theories, the Reisenberger model, and the so-called *interpolating model* (defined by Oeckl). The latter is given by a one-parameter family of models that interpolate between the trivial BF topological model and the Barrett–Crane model according to the value of a 'coupling constant'. Qualitative aspects of the renormalization groupoid flow of the couplings are studied in the various models.

Part II: The Barrett–Crane model

7. *SO*(4) Plebanski's action and the Barrett–Crane model

The Barrett–Crane model is one of the most extensively studied spin foam models for quantum gravity. In this section, we concentrate on the definition of the model in the Riemannian sector and sketch its generalization to the Lorentzian sector.

The Barrett–Crane model can be viewed as a spin foam quantization of the *SO*(4) Plebanski's formulation of general relativity. The general idea is analogous to that implemented in the model of section 5.1. We introduce the model from this perspective in the Riemannian sector so we need to start with the study of *Spin*(4) BF theory.

7.1. Quantum Spin(4) BF theory

Classical (*Spin*(4)) BF theory is defined by the action

$$S[B, A] = \int_{\mathcal{M}} \mathrm{Tr}[B \wedge F(A)], \tag{60}$$

where B^{IJ}_{ab} is a *Spin*(4) Lie-algebra-valued 2-form, A^{IJ}_{a} is a connection on a $Spin(4)$ principal bundle over $\mathcal{M}$. The theory has no local excitations. Its properties are very much analogous to the case of three-dimensional gravity studied in section 4.

[26] This interpretation becomes completely precise in the GFT formulation of spin foams presented in section 9.

A discretization Δ of $\mathcal{M}$ can be introduced along the same lines presented in section 4. For simplicity we concentrate on the case when Δ is a triangulation. The field B is associated with Lie algebra elements B_t assigned to triangles $t \in \Delta$. In four dimensions triangles $t \in \Delta$ are dual to faces $f \in \mathcal{J}_\Delta$. This one-to-one correspondence allows us to denote the discrete B by either a face (B_f) or a triangle (B_t) subindex, respectively. B_f can be interpreted as the 'smearing' of the continuous 2-form B on triangles in Δ. The connection A is discretized by the assignment of group elements $g_e \in Spin(4)$ to edges $e \in \mathcal{J}_\Delta$. The path integral becomes

$$Z(\Delta) = \int \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \prod_{f \in \mathcal{J}_\Delta} \mathrm{d}B_f\, \mathrm{e}^{\mathrm{i}B_f U_f} = \int \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \prod_{f \in \mathcal{J}_\Delta} \delta\left(g_{e_1} \cdots g_{e_n}\right), \quad (61)$$

where the first equality is the analogue of (29) while the second is the result of the B integration [99]. Using the Peter–Weyl theorem as in (32) one obtains

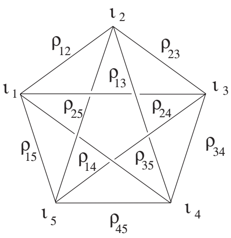
$$\mathcal{Z}(\Delta) = \sum_{\mathcal{C}:\{\rho\}\to\{f\}} \int \prod_{e \in \mathcal{J}_\Delta} \mathrm{d}g_e \prod_{f \in \mathcal{J}_\Delta} \Delta_{\rho_f} \mathrm{Tr}\left[\rho_f\left(g_e^1 \cdots g_e^N\right)\right], \quad (62)$$

where ρ are $Spin(4)$ irreducible representations. Integration over the connection can be performed as in the case of three-dimensional gravity. In a triangulation Δ, the edges $e \in \mathcal{J}_\Delta$ bound precisely four different faces; therefore, the g_e in (62) appear in four different traces. The relevant formula is

$$P^4_{\mathrm{inv}} := \int \mathrm{d}g\, \rho_1(g) \otimes \rho_2(g) \otimes \cdots \otimes \rho_4(g) = \sum_\iota C^\iota_{\rho_1\rho_2\cdots\rho_4} C^{*\iota}_{\rho_1\rho_2\cdots\rho_4}, \quad (63)$$

where P^4_{inv} is the projector onto $\mathrm{Inv}[\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n]$ and on the RHS we have expressed the projector in terms of normalized intertwiners as in (33). Finally, integration over the connection yields

$$Z_{\mathrm{BF}}(\Delta) = \sum_{\mathcal{C}_f:\{f\}\to\rho_f} \sum_{\mathcal{C}_e:\{e\}\to\{\iota_e\}} \prod_{f \in \mathcal{J}_\Delta} \Delta_\rho \prod_{v \in \mathcal{J}_\Delta}$$



, (64)

where the pentagonal diagram representing the vertex amplitude denotes the trace of the product of five normalized 4-intertwiners $C^\iota_{\rho_1\rho_2\rho_3\rho_4}$. As in the model for three-dimensional gravity, the vertex amplitude corresponds to the flat evaluation of the spin network state defined by the pentagonal diagram in (64), a $15j$-symbol. Vertices $v \in \mathcal{J}_\Delta$ are in one-to-one correspondence with 4-simplexes in the triangulation Δ. In addition we have $\mathcal{C}_e : \{e\} \to \{\iota_e\}$ representing the assignment of intertwiners to edges coming from (63).

The state sum (64) is generically divergent (due to the gauge freedom analogous to (24)). A regularized version defined in terms of $SU_q(2) \times SU_q(2)$ was introduced by Crane and Yetter [52, 100]. As in three dimensions (64) is topologically invariant and the spin foam path integral is discretization independent.

7.2. Classical SO(4) Plebanski action

Plebanski's Riemannian action can be thought of as *Spin*(4) BF theory plus constraints on the B field. It depends on an $so(4)$ connection A, a Lie-algebra-valued 2-form B and Lagrange multiplier fields λ and μ. Writing explicitly the Lie-algebra indices, the action is given by

$$S[B, A, \lambda, \mu] = \int \left[B^{IJ} \wedge F_{IJ}(A) + \lambda_{IJKL} B^{IJ} \wedge B^{KL} + \mu \epsilon^{IJKL} \lambda_{IJKL}\right], \tag{65}$$

where μ is a 4-form and $\lambda_{IJKL} = -\lambda_{JIKL} = -\lambda_{IJLK} = \lambda_{KLIJ}$ is a tensor in the internal space. Variation with respect to μ imposes the constraint $\epsilon^{IJKL}\lambda_{IJKL} = 0$ on λ_{IJKL}. The Lagrange multiplier tensor λ_{IJKL} then has 20 independent components. Variation with respect to λ imposes 20 algebraic equations on the 36 components of B. Solving for μ they are

$$\epsilon^{\mu\nu\rho\sigma} B^{IJ}_{\mu\nu} B^{KL}_{\rho\sigma} = e\epsilon^{IJKL}, \tag{66}$$

which is equivalent to

$$\epsilon_{IJKL} B^{IJ}_{\mu\nu} B^{KL}_{\rho\sigma} = e\epsilon_{\mu\nu\rho\sigma}, \tag{67}$$

for $e \neq 0$ where $e = \frac{1}{4!}\epsilon_{OPQR} B^{OP}_{\mu\nu} B^{QR}_{\rho\sigma}\epsilon^{\mu\nu\rho\sigma}$ [101]. The solutions to these equations are

$$B = \pm{}^*(e \wedge e), \qquad \text{and} \qquad B = \pm e \wedge e, \tag{68}$$

in terms of the remaining 16 degrees of freedom of the tetrad field e^I_a. If one substitutes the first solution into the original action one obtains Palatini's formulation of general relativity

$$S[e, A] = \int \mathrm{Tr}[e \wedge e \wedge {}^*F(A)]. \tag{69}$$

This property is the key to the definition of the spin foam model for gravity of the following section.

7.3. Discretized Plebanski's constraints

In section 7.1, we have derived the spin foam model for *Spin*(4) BF theory. We have seen that the sum over B-configurations is encoded in the sum over unitary irreducible representations of *Spin*(4). Can we restrict the spin foam configurations to those that satisfy Plebanski's constraints? To answer this question one first needs to translate the constraints of Plebanski's formulation into the simplicial framework. There are two ways in which we can approach the problem. One is by looking at the constraints in the form (65) and the other is by using (67).

The first version of the constraints is of the type of Ω^{ij} of section 5.1. This possibility has not been explored in the literature. A simple representation of the constraints is obtained if we choose to discretize (67), for which there are no free algebra indices.

As explained in section 7.1 in the discretization of the B field we replace the spacetime indices of the 2-form by a triangle:

$$\mu\nu \to t = \text{triangle}. \tag{70}$$

Using this it is easy to translate (67) into the simplicial framework. The RHS of (67) vanishes when there are repeated spacetime indices. At the discrete level, this means that whenever the triangles t and t' belong to the same tetrahedron T, $t, t' \in T$, (i.e., either $t = t'$ or t and t' share an edge) we have

$$\epsilon_{IJKL} B^{IJ}_t B^{KL}_{t'} = 0. \tag{71}$$

The remaining constraints correspond to the situation when t and t' do not lie on the same tetrahedron (no 'repeated spacetime indices'). We can write these constraints as follows. If we label a vertex in a given 4-simplex, $0, \ldots, 4$ we can label the six triangles $0jk$ containing the vertex 0 simply as jk $(j, k = 1, \ldots, 4)$. Only these are needed to express the independent constraints

$$\epsilon_{IJKL} B^{IJ}_{12} B^{KL}_{34} = \epsilon_{IJKL} B^{IJ}_{13} B^{KL}_{42} = \epsilon_{IJKL} B^{IJ}_{14} B^{KL}_{23} \propto e. \tag{72}$$

Now we want to implement the constraints. The strategy is to impose them directly on the BF spin foam sum, (62), after the B integration has been performed. Formally one would associate the discrete B_f with the differential operator $-\mathrm{i}\partial/\partial U_f$ in (61). More precisely, the observation is that the *Spin*(4) left-invariant vector field $-\mathrm{i}\mathcal{X}^{IJ}(U) := U^{\mu}{}_{\nu}X^{IJ\nu}{}_{\sigma}\frac{\partial}{\partial U^{\mu}{}_{\sigma}}$ acts as a quantum B^{IJ} on (61) since

$$-\mathrm{i}\mathcal{X}^{IJ}(U)(\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]})|_{U\sim 1} = U^{\mu}{}_{\nu}X^{IJ\nu}{}_{\sigma}\frac{\partial}{\partial U^{\mu}_{\sigma}}\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}|_{U\sim 1}$$
$$= \mathrm{Tr}[UX^{IJ}B]\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}|_{U\sim 1} \sim B^{IJ}\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}, \tag{73}$$

where X^{IJ} are elements of an orthonormal basis in the *SO*(4) Lie algebra. The evaluation at $U = 1$ is motivated by the fact that configurations in the BF partition function (61) have support on flat connections. This approximation is made in order to motivate our definition but it plays no role in the implementation of the constraints.

The constraints (67) are quadratic in the *B*. We have to worry about cross terms; the non-trivial case corresponds to

$$\epsilon_{IJKL}\mathcal{X}^{IJ}(U)\mathcal{X}^{KL}(U)(\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]})|_{U\sim 1}$$
$$= -\epsilon_{IJKL}\big(\mathrm{Tr}[X^{IJ}UB]\mathrm{Tr}[X^{KL}UB]\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]} + \mathrm{i}\mathrm{Tr}[X^{IJ}X^{KL}UB]\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}\big)|_{U\sim 1}$$
$$\sim \epsilon_{IJKL}B^{IJ}B^{KL}\,\mathrm{e}^{\mathrm{i}\mathrm{Tr}[BU]}, \tag{74}$$

where the second term on the second line can be dropped using that $\epsilon_{IJKL}X^{IJ}X^{KL} \propto 1$ (one of the two *SO*(4) Casimir operators) and $U \sim 1$. Therefore, we define the B_f field associated with a face at the level of equation (32) as the appropriate left-invariant vector field $-\mathrm{i}\mathcal{X}^{IJ}(U_f)$ acting on the corresponding discrete holonomy U_f, namely

$$B_f^{IJ} \to -\mathrm{i}\mathcal{X}^{IJ}(U_f). \tag{75}$$

Gauge invariance of the BF partition function implies that for every tetrahedron

$$\sum_{t\in T} B_t^{IJ} = 0, \tag{76}$$

where $t \in T$ denotes the triangles in the corresponding tetrahedron.

In order to implement (71) we concentrate on a single 4-simplex amplitude, i.e., the fundamental *atom* of figure 4. Once the constrained 4-simplex amplitude is constructed any spin foam amplitude can be obtained by gluing atoms together along faces by integration over common boundary data as in section 3.2. The BF 4-simplex wavefunction is obtained using (61) on the dual 2-complex with the boundary defined by the intersection of the dual of a single 4-simplex with a 3-sphere, see figure 4.

The amplitude of the fundamental *atom* is a (cylindrical) function depending on the boundary values of the connection on the boundary graph γ_5. We denote as $h_{ij} \in Spin(4)$ ($i \neq j, i, j = 1, \ldots, 5$ and $h_{ij} = h_{ji}^{-1}$) the corresponding ten boundary variables (associated with thin boundary edges in figure 4)[27] and $g_i \in Spin(4)$ ($i = 1, \ldots, 5$) the internal connection (corresponding to the thick edges in figure 4). According to (61) the 4-simplex BF amplitude $4SIM_{\mathrm{BF}}(h_{ij})$ is given by

$$4SIM_{\mathrm{BF}}(h_{ij}) = \int \prod_i \mathrm{d}g_i \prod_{i<j} \delta(g_i h_{ij} g_j), \tag{77}$$

[27] Strictly speaking, the boundary connections h_{ij} are defined as the product $h'_{ij}h''_{ij}$, where h' and h'' are associated with half-paths as follows: take the edge ij for simplicity and assume it is oriented from i to j. Then h'_{ij} is the discrete holonomy from i to some point in the centre of the path and h''_{ij} is the holonomy from that centre point to j. This splitting of variables is necessary when matching different atoms to reconstruct the simplicial amplitude [56, 57].

where $U_{ij} = g_i h_{ij} g_j$ is the holonomy around the triangular face $0ij$. With the definition of the B fields given in (75) the constrained amplitude, $4SIM_{\text{const}}(h_{ij})$, formally becomes

$$4SIM_{\text{const}}(h_{ij}) = \int \prod_i \mathrm{d}g_i \delta[\text{Constraints}(-\mathrm{i}\mathcal{X}(U_{ij}))] \prod_{i<j} \delta(g_i h_{ij} g_j). \qquad (78)$$

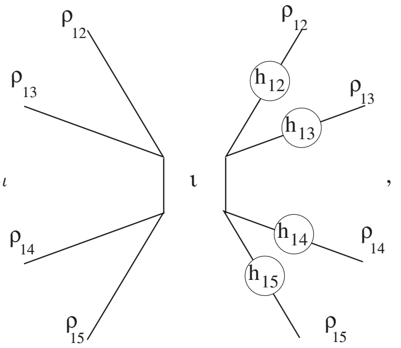
It is easy to verify, using an equation analogous to (73) and the invariance of ϵ_{IJKL}, that one can define the B by simply acting with the left-invariant vector fields on the boundary connection h_{ij}. Therefore, the previous equation is equivalent to

$$4SIM_{\text{const}}(h_{ij}) = \delta[\text{Constraints}(-\mathrm{i}\mathcal{X}(h_{ij}))] \int \prod_i \mathrm{d}g_i \prod_{i<j} \delta(g_i h_{ij} g_j), \qquad (79)$$

where we have taken the delta function out of the integral. The quantity on which the formal delta distribution acts is simply $4SIM_{\text{BF}}(h_{ij})$. The amplitude $4SIM_{\text{BF}}(h_{ij})$ can be expressed in a more convenient way if we expand the delta functions in modes as in (62) and then integrate over the internal connection g_i. The integration is analogous to that in (63), for example integration over g_1 yields

$$P^4_{\text{inv}} \rho_{12}(h_{12}) \otimes \rho_{13}(h_{13}) \otimes \rho_{14}(h_{14}) \otimes \rho_{15}(h_{15})$$

$$= \sum_\iota \Delta_\iota$$

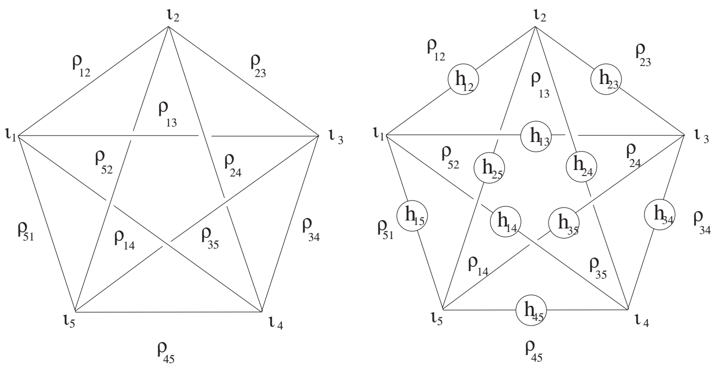


, (80)

where on the RHS we have chosen a particular basis of 4-intertwiners to span the projector P^4_{inv} into $\text{Inv}[\rho_{12}\cdots\rho_{15}]$. The 4-intertwiners are explicitly given as contractions of normalized intertwiners using the graphical notation introduced in section 4. The circles represent the corresponding ρ-representation matrices evaluated on the corresponding boundary connection h. The 4-simplex amplitude becomes

$$4SIM_{\text{BF}}(h_{ij})$$

$$= \sum_{\rho_{ij},\iota_i} \prod_{i<j} \Delta_{\rho_{ij}}$$



. (81)

In the previous equation four-valent nodes denote normalized 4-intertwiners and the tree decomposition is left implicit (the factors Δ_ι in (80) have been absorbed into the notation).

The term on the LHS is a $15j$-symbol as in (64) while the term on the RHS is the trace of five 4-intertwiners with the respective boundary connection insertions.

The 4-simplex amplitude for the constrained spin foam model is then defined as the restriction of $4SIM_{\mathrm{BF}}(h_{ij})$ imposed by the quantum version of the constraints (71) defined by the following set of differential equations:

$$\hat{C}_{ij,ik}4SIM_{\mathrm{const}}(h_{ij}) = \epsilon_{IJKL}\mathcal{X}^{IJ}(h_{ij})\mathcal{X}^{KL}(h_{ik})4SIM_{\mathrm{const}}(h_{ij}) = 0 \qquad \forall j, k, \tag{82}$$

and where the index $i = 1, \ldots, 5$ is held fixed and $\hat{C}_{ij,ik}$ denotes the corresponding constraint operator.

There are six independent constraints (82) for each value of $i = 1, \ldots, 5$. If we consider all the equations for the 4-simplex amplitude then some of them are redundant due to (76). The total number of independent conditions is 20, in agreement with the number of classical constraints (67). For a given i in (82) (i.e., a given tetrahedron) and for $j = k$ the equation becomes

$$\begin{aligned}
&\epsilon_{IJKL}\mathcal{X}^{IJ}(h_{ij})\mathcal{X}^{KL}(h_{ij})4SIM_{\mathrm{const}}(h_{ij}) \\
&\quad = \delta_{ik}\left[J^i\left(h^{\ell}_{ij}\right)J^k\left(h^{\ell}_{ij}\right) - J^i\left(h^{r}_{ij}\right)J^k\left(h^{r}_{ij}\right)\right]4SIM_{\mathrm{const}}\left(h^{\ell}_{ij}, h^{r}_{ij}\right) \\
&\quad = \left[j^{\ell}_{ij}\left(j^{\ell}_{ij}+1\right) - j^{r}_{ij}\left(j^{r}_{ij}+1\right)\right]4SIM_{\mathrm{const}}(h_{ij}) = 0,
\end{aligned} \tag{83}$$

where we have used that $Spin(4) = SU(2) \times SU(2)$ so that for $h \in Spin(4)$, $h^{\ell}, h^{r} \in SU(2)$ denote its right and left components, irreducible representation can be expressed as $\rho = j^{\ell} \otimes j^{r}$ for $j^{\ell}, j^{r} \in \mathrm{Irrep}[SU(2)]$, and the left-invariant vector field

$$\mathcal{X}^{IJ}(h) = P^{+IJ}_{i}J^{i}(h^{\ell}) + P^{-IJ}_{i}J^{i}(h^{r}), \tag{84}$$

for $\epsilon_{IJ}{}^{KL}P^{\pm IJ} = \pm P^{\pm KL}$ and J^i being the left-invariant vector fields on the corresponding left- and right-$SU(2)$ copies of $Spin(4)$. The previous constraints are solved by requiring the corresponding representation ρ_{ij} to be simple, i.e.,

$$\rho_{ij} = j_{ij} \otimes j^{*}_{ij} \qquad \text{or} \qquad \rho_{ij} = j_{ij} \otimes j_{ij}. \tag{85}$$

This ambiguity is analogous to the classical one in (68). We take $\rho_{ij} = j_{ij} \otimes j^{*}_{ij}$ in correspondence with the choice ${}^{*}(e \wedge e)$ that produces the gravity sector[28]. This solves 10 of the 20 equations. The next non-trivial conditions imposed by (82) is when $j \neq k$. In this case we have

$$\begin{aligned}
&2\epsilon_{IJKL}\mathcal{X}^{IJ}(h_{ij})\mathcal{X}^{KL}(h_{ik})4SIM_{\mathrm{const}}(h_{ij}) \\
&\quad = \epsilon_{IJKL}(\mathcal{X}^{IJ}(h_{ij}) + \mathcal{X}^{IJ}(h_{ik}))(\mathcal{X}^{KL}(h_{ij}) + \mathcal{X}^{KL}(h_{ik}))4SIM_{\mathrm{const}}(h_{ij}) \\
&\quad = [\iota^{\ell}(\iota^{\ell}+1) - \iota^{r}(\iota^{r}+1)]4SIM_{\mathrm{const}}(h_{ij}) \\
&\quad = 0,
\end{aligned} \tag{86}$$

where in the second line we used the fact that we have already solved (83). In the third line we have used the gauge invariance (or the analogue of (76) for the three-valent node in the tree decomposition that pairs the representation ρ_{ij} with the ρ_{ik}) in order to express the sum of invariant vector fields as the invariant vector field acting on the virtual link labelled by ι. This choice of tree decomposition in the case $ij = 12$ and $ik = 13$ is that used in equation (80). The solution is clearly $\iota = \iota \otimes \iota^{*}$.

What happens now to either of the two remaining conditions (only one is independent), for example, $\hat{C}_{ij',ik'}$ for $k \neq k'$, $j \neq j'$ and $j' \neq k'$? At first sight it looks like these

[28] In a detailed analysis Baez and Barrett [15] justify this choice in a rigorous way. This restriction (imposed by the so-called chirality constraint, section 7.7) implies that the 'fake tetrahedron' configurations, corresponding to the solutions of the constraints on the right-hand side of equation (68), are dropped from the state sum.

equations cannot be (generically) satisfied because an intertwiner that has simple ι in one tree decomposition does not have only simple ι' components in a different tree decomposition as a consequence of the recoupling identity (42). There is, however, a linear combination of intertwiners found by Barrett and Crane in [9] which is simple in any tree decomposition, namely

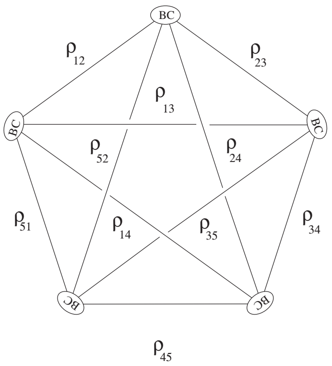
$$|\Psi_{\mathrm{BC}}\rangle = \sum_{\mathrm{simple}\,\iota} C^{\iota}_{\rho_1,\ldots,\rho_4}. \tag{87}$$

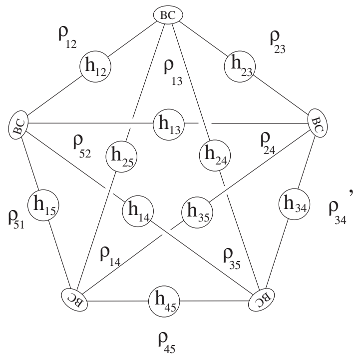
$C^{\iota}_{\rho_1,\ldots,\rho_4}$ is a normalized 4-intertwiner and the summation is over simple ι (i.e. $\iota = \iota \otimes \iota^*$) and the ρ_i are also simple ($\rho_i = j_i \otimes j_i^*$ for $i = 1, \ldots, 4$). This is clearly a solution to all the constraints and has been shown to be the unique one (up to an overall factor) by Reisenberger in [102]. $|\Psi_{\mathrm{BC}}\rangle$ defines the so-called Barrett–Crane intertwiner. Now the projector P^4_{inv} in (80), the building block of the BF amplitude, can be written as

$$P^4_{\mathrm{inv}} = |\Psi_{\mathrm{BC}}\rangle\langle\Psi_{\mathrm{BC}}| + \text{orthogonal terms}, \tag{88}$$

using the standard Gram–Schmidt construction of a basis in $\mathrm{Inv}[\rho_1 \cdots \rho_4]$. In other words P^4_{inv} is the sum of one-dimensional projector to the solutions of the constraints (82) plus the orthogonal complement. The solution to (82) is then unique (up to scaling) and can be written as

$4SIM_{\mathrm{const}}(h_{ij})$

$$= \sum_{j_{ij},\iota_i} \prod_{i<j} \Delta_{\rho_{ij}}$$ 

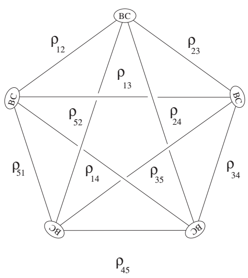


, (89)

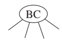
where the $\rho_{ij} = j_{ij} \otimes j^*_{ij}$ and we graphically represent $|\Psi_{BC}\rangle$ by 

.

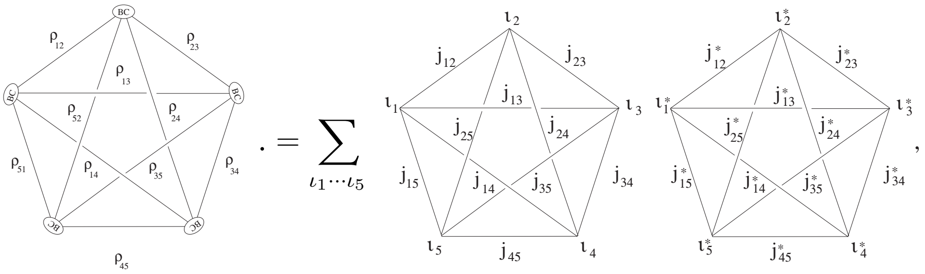
The amplitude of an arbitrary simplicial complex is computed by putting together 4-simplexes with consistent boundary connections and gluing them by integration over boundary data. If we do that we obtain the Barrett–Crane state sum on a fixed triangulation

$$Z_{\mathrm{BC}}(\mathcal{J}_\Delta) = \sum_{\mathcal{C}_f:\{f\}\to j_f} \prod_{f\in\mathcal{J}_\Delta} (2j_f+1)^2 \prod_{e\in\mathcal{J}_\Delta} A_e \prod_{v\in\mathcal{J}_\Delta}$$ 

, (90)

where we have made the replacement $\Delta_\rho = (2j+1)^2$ for $\rho = j \otimes j^*$. The vertex amplitude depends on the ten representations labelling the ten faces in a 4-simplex and it is referred to as the $10j$-symbol. Using the definition (87) of the Barrett–Crane intertwiner, 

, the $10j$-symbol can be written explicitly in terms of $15j$-symbols as



(91)

where we represent the normalized $SO(4)$ intertwiners $C^{\iota}_{\rho_1\cdots\rho_4}$ in (87) in terms of the corresponding pair of left–right $SU(2)$ intertwiners.

A_e denotes a possible edge amplitude which is not determined by our argument. The implementation of constraints in the path integral should be supplemented with the appropriate modification of the measure. This should affect the values of lower-dimensional simplexes such as face and edge amplitudes. Constraints (82) act on each edge (tetrahedron) separately; heuristically one would expect a Jacobian factor to modify and so determine the value of the edge amplitude A_e. As we pointed out in section 6.1, the resolution of this ambiguity is related to the question of how to define the correct anomaly-free measure. In section 9, we will see how the GFT formulation provides a natural definition of A_e.

Constraints that involve different tetrahedra in a given 4-simplex, corresponding to (72), are automatically satisfied as operator equations on the Barrett–Crane solutions. This can be checked using (82) and (76).

7.4. *The quantum tetrahedron in 4D and the Barrett–Crane intertwiner*

We have re-derived the Barrett–Crane model directly from a simplicial formulation of Plebanski's action. The original definition makes use of the concept of the 'quantum tetrahedron' [14, 15]. In that context, the analogue of our B is given by the bivectors associated with the triangles of a classical tetrahedron. A Hilbert space is defined using geometric quantization and the classical triangle bivectors are promoted to operators. This Hilbert space is reduced by implementation of constraints—quantization of the geometric constraints satisfied by the classical tetrahedron—to the Hilbert space of the so-called quantum tetrahedron. These constraints are precisely of the form (72). For fixed triangle quantum numbers the state of the quantum tetrahedron is defined by the Barrett–Crane intertwiner (87). There is, however, no systematic prescription for the construction of the state sum and even the single 4-simplex amplitude has to be given as an, although natural, *ad hoc* definition.

As we have shown in the previous sections, all these can be obtained from the path integral approach applied to (65). In addition to simplifying the derivation—no additional quantization principle is required—our framework provides a direct and systematic definition of the state sum. In [14, 15] one can only access the constraint operator algebra of a single tetrahedron and conclude that this sub-algebra is closed. The full constraint algebra of (82) can only be studied in the formalism presented here. Finally, since our prescription is directly derived from an action principle it is conceivable that a rigorous derivation of the undetermined amplitudes (such as A_e in (90)) could exist. These important questions remain to be investigated.

7.5. An integral expression for the 10j-symbol

In section 9, we will present a discretization-independent formulation of the Barrett–Crane model based on a GFT. The realization by Barrett [103] that the vertex amplitude (91) admits an integral representation plays a key role in the construction. The integral formula is also important in the computation of the $10j$-symbol asymptotics that we briefly describe in the following subsection.

The basic observation in the construction of an integral expression for (91) is that equation (87) has precisely the form (33) (for $n = 4$) if we write the *Spin*(4) intertwiners as tensor products of $SU(2)$ ones. Therefore,

$$|\Psi_{\mathrm{BC}}\rangle = \int_{SU(2)} \mathrm{d}u\, j_1(u) \otimes j_2(u) \otimes j_3(u) \otimes j_4(u), \tag{92}$$

where $j(u)$ denotes $SU(2)$ representation matrices in the representation j. Each one of the five pairs of intertwiners in (91) can be obtained as an integral (92) over $SU(2)$. Each of the ten representation matrices j_{ik} $(i \neq k = 1, \ldots, 5)$ appears in two integrals corresponding to the intertwiners at the nodes i and j, respectively. Contracting the matrix indices according to (91) these two representation matrices combine into a trace $\mathrm{Tr}\big[j_{ik}\big(u_i u_k^{-1}\big)\big]$ $(u_i \in SU(2))$. Parametrizing $SU(2)$ with spherical coordinates on S^3

$$\mathrm{Tr}\left[j_{ik}\left(u_i u_k^{-1}\right)\right] = \frac{\sin(2j_{ik}+1)\psi_{ik}}{\sin(\psi_{ik})} := (2j_{ik}+1) K_{j_{ik}}(y_i, y_k), \tag{93}$$

where ψ_{ik} is the azimuthal angle between the points y_i, y_k on the sphere corresponding to u_i and u_k, respectively. We have also introduced the definition of the kernel $K_{j_{ik}}(y_i, y_k)$ in terms of which the Barret–Crane vertex amplitude (91) becomes

$$A_v(j_{ik}) = \int_{(S^3)^5} \prod_{i=1}^{5} \mathrm{d}y_i \prod_{i<k} (2j_{ik}+1) K_{j_{ik}}(y_i, y_k). \tag{94}$$

Each of the five integration variables in S^3 can be regarded as a unit vector in R^4. They are interpreted as unit normal vectors to the three-dimensional hyperplanes spanned by the corresponding five tetrahedra. The angle ψ_{ik} is defined by $\cos\psi_{ik} = y_i \cdot y_k$ and corresponds to the exterior angle between two hyperplanes (analogous to the dihedral angles of Regge calculus). These normals determine a 4-simplex in R^4 up to translations and scaling [103].

7.6. The asymptotics for the vertex amplitude

The large-spin behaviour of the spin foam amplitudes provides information about the low-'energy' or semiclassical limit of the model [104] in the naive sense $\hbar \to \infty$ while geometric quantities such as area (7) are held fixed. Evidence showing a connection between the asymptotics of the Barrett–Crane vertex and the action of general relativity was found by Crane and Yetter in [105].

A computation of the asymptotic (large-j) expression of the Barrett–Crane vertex amplitude for non-degenerate configurations was obtained by Barrett and Williams in [106]. They computed $A_v(j_{ik})$ for large j_{ik} by looking at the stationary phase approximation of the oscillatory integral (94). The large-spin behaviour of the vertex amplitude is given by

$$A_v(j_{ik}) \sim \sum_{\sigma} P(\sigma) \cos\left[S_{\mathrm{Regge}}(\sigma) + \kappa \frac{\pi}{4}\right] + D, \tag{95}$$

where the sum is over geometric 4-simplexes σ whose face areas are fixed by the spins. The action in the argument of the cosine corresponds to Regge action which in four dimensions is defined by $S_{\mathrm{Regge}}(\sigma) = \sum_{i<k} A_{ik}\,\psi_{ik}(\sigma)$ where A_{ik} is the area of the ik-triangle. $P(\sigma)$ is a normalization factor which does not oscillate with the spins. D is the contribution of degenerate configurations, i.e. those for which some of the hyperplane normals defined above coincide. However, in a recent paper [107] Baez *et al* show that the term D is in fact dominant in the previous equation, i.e., the leading-order terms are contained in the set of degenerate configurations! This was later confirmed by the results of Freidel and Louapre [108] and Barrett and Steele [109].

7.7. Area and volume in the Barrett–Crane model

Using the representation of the discrete B of section 7.3, we can define the operator corresponding to the square of B_t corresponding to any triangle $t \in \Delta$, namely

$$\hat{B}_{tIJ}\hat{B}_t^{IJ} = -\mathcal{X}_{IJ}(U_f)\mathcal{X}^{IJ}(U_f),$$

where f is the face dual to $t \in \Delta$. Acting on the modes of the expansion (62) this is simply the Casimir

$$\hat{B}_{tIJ}\hat{B}_t^{IJ} = [j^\ell(j^\ell+1) + j^r(j^r+1)]1. \tag{96}$$

It is easy to check that this operator commutes with all the constraints and therefore is well defined on the space of solutions of the constraints. Its geometrical meaning is clear if we recall that at the classical level B_t represents the bivector associated with the triangle $t \in \Delta$. Once the constraints are imposed (96) is proportional to the square of the area of the corresponding triangle. The simplicity constraint implies $j_\ell = j_r^* = j$ so that the areas of the triangles have discrete eigenvalues a_j given by

$$a_j \propto \sqrt{j(j+1)}, \tag{97}$$

in agreement with the result of quantum geometry (7).[29]

Can we define the analogue of the volume operator of quantum geometry? The candidate for the square of such an operator is

$$U_\pm = \epsilon_{ijk}\big[J^i_{R1}J^j_{R2}J^k_{R3} \pm J^i_{L1}J^j_{L2}J^k_{L3}\big], \tag{98}$$

where J_1, J_2 and J_3 are the operators defined in (84) corresponding to three different triangles in the tetrahedron. The operator U_+ vanishes identically on the solutions of the quantum constraints. In fact it can be expressed as the commutator of the constraints (82). This is the chirality constraint of [15].

One would like to define the volume operator as the square root of U_-; however, U_- is not a well-defined operator on the Hilbert space of the quantum tetrahedron since it does not commute with the simplicity constraints. In other words, the action of the volume operators maps states out of the solution of the constraints. Only when the dimension of $\mathrm{Inv}\,[j_1 \otimes j_2 \otimes j_3 \otimes j_4]$ is 1 does the commutator vanish (this is in fact a necessary and sufficient condition). The volume operator is well defined in this subspace but it vanishes. In conclusion, generically the volume operator is not defined in the Barrett–Crane model.

7.8. Lorentzian generalization

A Lorentzian generalization of the Riemannian Barrett–Crane vertex amplitude was proposed in [10]. Using the spin foam model GFT duality of section 8 a generalization of the full

[29] Given a triangulation Δ of $\mathcal{M}$ and the induced triangulation Δ_Σ of a slice $\Sigma \subset \mathcal{M}$ any spin foam defined on $\mathcal{J}_\Delta$ induces a spin network state on a graph γ_Σ dual to Δ_Σ. Links of γ_Σ are dual to triangles in Δ_Σ. These triangles play the role of the surface S in the definition of the area operator (7).

model (including face and edge amplitudes) was found in [13]. The relevant representations of the Lorentz group are the unitary ones. Unitary irreducible representations of $SL(2,\mathbb{C})$ are infinite dimensional and labelled by $n \in \mathbb{N}$ and $\rho \in \mathbb{R}^+$. The simplicity constraints select the representations for which $n = 0$. The triangle area spectrum is given by

$$a_\rho \propto \sqrt{\rho^2 + 1}. \tag{99}$$

There is still a minimum eigenvalue but the spectrum is continuous. The derivation of the model can be obtained from the $SO(3,1)$ Plebanski's action by generalizing the computations of section 7. We will study this model in section 9.3 where we show that the Lorentzian extension is not unique. In fact a new Lorentzian model can be defined following the Barrett–Crane prescription [110].

7.9. *Positivity of spin foam amplitudes*

Baez and Christensen [111] showed that spin foam amplitudes of the (Riemannian) Barrett–Crane model are positive for any closed 2-complex. For open spin foams they are real and (if not zero) their sign is given by $(-1)^{2J}$ where J is the sum of the spin labels of the edges of the boundary spin networks. This is a rather puzzling property since (as the authors point out) this seems to imply the absence of quantum interference. Positivity of the Lorentzian model seems to hold according to numerical evaluations of the vertex amplitude.

7.10. *Degenerate sectors of Spin(4) Plebanski's formulation*

In [56] Reisenberger showed that, solving the constraints (66) in the degenerated sector, i.e., where $e = \frac{1}{4!}\epsilon_{OPQR} B^{OP}_{\mu\nu} B^{QR}_{\rho\sigma} \epsilon^{\mu\nu\rho\sigma} = 0$, and substituting the solution back into the Plebanski action, the theory reduces to two sectors described by the actions

$$S^{\pm}_{\text{deg}} = \int B^r_i \wedge \big(F_i(A^r) \pm V^j_i F_j(A^\ell)\big), \tag{100}$$

where the upper index r (respectively ℓ) denotes the self-dual (respectively anti-self-dual) part of B and A in the internal space, and $V \in SO(3)$. Let us concentrate on the sector with the minus sign in the previous expression. Then it is straightforward to define the discretized path integral along the same lines as BF theory in section 7.1. The result is

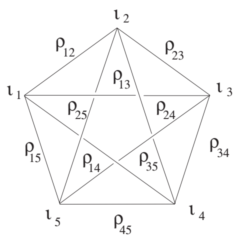
$$\mathcal{Z}(\Delta) = \int \prod_{f\in\Delta^*} \mathrm{d}B^{r(3)}_f \, \mathrm{d}v_f \prod_{e\in\Delta^*} \mathrm{d}g^\ell_e \, \mathrm{d}g^r_e \, \mathrm{e}^{\mathrm{i}\mathrm{Tr}[B^r_f U^r_f v_f U^{\ell-1}_f v^{-1}_f]}, \tag{101}$$

where $\mathrm{d}g^\ell_e$, $\mathrm{d}g^r_e$ and $\mathrm{d}v_f$ are defined in terms of the $SU(2)$ Haar measure. Integrating over the B field we obtain

$$Z(\Delta) = \int \prod_{e\in\Delta^*} \mathrm{d}g^\ell_e \, \mathrm{d}g^r_e \prod_{f\in\Delta^*} \mathrm{d}v_f \delta^{(3)}\big(g^r_{e_1}\cdots g^r_{e_n} v_f \big(g^\ell_{e_1}\cdots g^\ell_{e_n}\big)^{-1} v^{-1}_f\big), \tag{102}$$

where the delta function $\delta^{(3)}$ denotes an $SU(2)$ distribution.

Integration over g^ℓ_e, g^r_e and v_f yields a spin foam model where only face representations are constrained to be simple while intertwiners are arbitrary. Explicitly

$$Z_{\text{deg}}(\Delta) = \sum_{\mathcal{C}_f:\{f\}\to j_f} \sum_{\mathcal{C}_e:\{e\}\to\{\iota_e\}} \prod_{v\in\Delta^*}$$ 

, (103)

where $\rho_f = j_f \otimes j_f^*$ and $j_f \in \mathrm{Irrep}[SU(2)]$. This is precisely the spin foam obtained in [11]! This model was obtained as a natural modification of the GFT that defines a variant of the BC model. Here we have rediscovered the model from the simplicial quantization of S^-_{deg} defined in (100). This establishes the relation of the model with a classical action! It corresponds to spin foam quantization of the '−' degenerate sector of $SO(4)$ Plebanski's theory.

The + sector action (100) can be treated in a similar way. The only modification is that of the subgroup. Instead of using the diagonal insertion defined above one has to define $u \in SU(2) \subset Spin(4)$ so that $ug = (ug^{\ell}, u^{-1}g^{r})$. This selects representations of the form $\rho = j \otimes j$ instead of $\rho = j \otimes j^*$ for faces.

Note that the allowed 4-simplex configurations of the model of section 7.10 are fully contained in the set of 4-simplex configurations of the model obtained here.

The positivity results have been generalized to generic 2-complexes (not necessarily dual to a triangulation) by Pfeiffer [112], who also shows that the sign -1^{2J} is not really present if one chooses the correct bases of intertwiners.

8. The group field theory (GFT) ansatz

In this section, we present the motivation and main results of [16, 17] where the duality relation between spin foam models and group field theories is established. This is the formulation we referred to in section 6.2 as one of the possible discretization-independent definitions of spin foam models. The result is based on the generalization of matrix models introduced by Boulatov [86] and Ooguri [87] dual to BF theory in three and four dimensions, respectively. This formulation was first proposed for the Riemannian Barrett–Crane model in [11] and then generalized to a wide class of spin foam models [16, 17].

Given a spin foam model defined on an arbitrary 2-complex $\mathcal{J}_\Delta$ (dual to a triangulation Δ)—thus the partition function $Z[\mathcal{J}_\Delta]$ is of the form (21)—there exists a GFT such that the perturbative expansion of the field theory partition function generalizes (21) to a sum over 2-complexes represented by the Feynman diagrams $\mathcal{J}$ of the field theory. These diagrams look 'locally' as dual to triangulations (vertices are five-valent, edges are four-valent) but they are no longer tied to any manifold structure [88].

We now motivate the duality using what we know of the model defined on a fixed simplicial decomposition. The action of the GFT is of the form

$$I[\phi] = I_0[\phi] + \frac{\lambda}{5!} V[\phi], \qquad (104)$$

where $I_0[\phi]$ is the 'kinetic' term quadratic in the field and $V[\phi]$ denotes the interaction term. The field ϕ is defined below. The expansion in λ of the partition function takes the form

$$\mathcal{Z} = \int \mathcal{D}[\phi]\, \mathrm{e}^{-I[\phi]} = \sum_{\mathcal{J}_N} \frac{\lambda^N}{\mathrm{sym}(\mathcal{J}_\mathrm{N})} Z[\mathcal{J}_N], \qquad (105)$$

where $\mathcal{J}_N$ is a Feynman diagram (2-complex) with N vertices and $Z[\mathcal{J}_N]$ is the one given by (21). In fact, the interaction term $V[\phi]$ is fixed uniquely by the 4-simplex amplitude of the simplicial model while the kinetic term $I_0[\phi]$ is trivial as we argue below.

Expression (105) is taken as the discretization- (and manifold-) independent definition of the model. Transition amplitudes between spin network states on boundary graphs γ_1 and γ_2 are shown to be given by the correlation functions

$$\int \mathcal{D}[\phi] \underbrace{\phi \cdots \phi}_{\gamma_1} \underbrace{\phi \cdots \phi}_{\gamma_2} \mathrm{e}^{-I[\phi]}, \qquad (106)$$

where boundary graphs are determined by the arrangement of fields in the product. The field can be defined as operators creating four-valent nodes of spin networks [92, 113]. In [92] correlation functions $\langle\phi\cdots\phi\rangle$ are interpreted as a complete family of gauge-invariant observables for quantum gravity. They encode (in principle) the physical content of the theory and can be used to reconstruct the physical Hilbert space in a way that mimics Wightman's procedure for standard QFT [114, 115].

How do we construct the GFT out of the spin foam model on a fixed 2-complex? In (105) the combinatoric of diagrams is completely fixed by the form of the action (104). To construct the action of the GFT one starts from the spin foam model defined on a 2-complex $\mathcal{J}_\Delta$ dual to a simplicial complex Δ. The sum over 2-complexes in (106) contains only those 2-complexes that locally look like the dual of a simplicial decomposition.

Spin foams on such 2-complexes have edges *e* with which are associated tensor products of four representations $\rho_1\otimes\rho_2\otimes\rho_3\otimes\rho_4$, since they bound four coloured faces in $\mathcal{J}_\Delta$. If we want to think of the edge *e* as associated with the propagator of a field theory then such a propagator should be a map

$$\mathcal{P}:\rho_1\otimes\rho_2\otimes\rho_3\otimes\rho_4\rightarrow\rho_1\otimes\rho_2\otimes\rho_3\otimes\rho_4. \tag{107}$$

According to the Peter–Weyl theorem, elements of $\rho_1\otimes\cdots\otimes\rho_4$ naturally appear in the mode expansion of a function $\phi(x_1,\ldots,x_4)$ for $x_i\in G$. Spin labels arise as 'momentum' variables in the field theory. Intertwiners assigned to edges in spin foams impose compatibility conditions on the representations. In the context of the GFT this is interpreted as 'momentum conservation' which is guaranteed by the requirement that the field be 'translational invariant',

$$\phi(x_1,x_2,x_3,x_4)=\phi(x_1g,x_2g,x_3g,x_4g)\qquad\forall g\in G. \tag{108}$$

One also requires ϕ to be invariant under permutations of its arguments. One can equivalently take the field ϕ to be arbitrary and impose translation invariance by acting with $\mathbf{P}$ defined as

$$\mathbf{P}\phi(x_1,x_2,x_3,x_4)=\int\mathrm{d}g\,\phi(x_1g,x_2g,x_3g,x_4g). \tag{109}$$

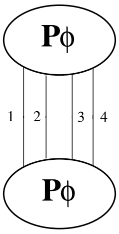
All the information on local spin foams is in the vertex amplitude or fundamental atom (4-simplex) amplitude of section 3.2. Therefore, the kinetic term $I_0[\phi]$ is simply given by

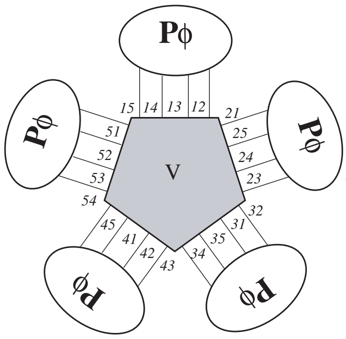
$$I_0[\phi]=\frac{1}{2}\int\mathrm{d}^4x[\mathbf{P}\phi(x_1,x_2,x_3,x_4)]^2. \tag{110}$$

Since vertices are five-valent in our discretization the interaction term should contain the product of five field operators. This is a function of 20 group elements. If we use the compact notation $\phi(x_i):=\phi(x_1,\ldots,x_4)$, then the general 'translation-invariant' form is

$$V[\phi]=\int\mathrm{d}^{20}x\,\mathcal{V}(x_{ij}[x_{ji}]^{-1})\mathbf{P}\phi(x_{1i})\mathbf{P}\phi(x_{2i})\mathbf{P}\phi(x_{3i})\mathbf{P}\phi(x_{4i})\mathbf{P}\phi(x_{5i}), \tag{111}$$

where $\mathcal{V}$ is a function of ten variables evaluated on the 'translation-invariant' combinations $\alpha_{ij}:=x_{ij}[x_{ji}]^{-1}$. For local spin foams the function $\mathcal{V}(\alpha_{ij})$ is in one-to-one correspondence with the fundamental 4-simplex (atom) amplitude (19). If we represent the field $\phi(x_1, x_2, x_3, x_4)$ by  we can write the action (104), using (110) and (111), as

$S[\phi] = \int d^4x$ 

$+ \frac{\lambda}{5!} \int d^{20}x$ 

, (112)

where the straight lines represent the field arguments (in the case of the interaction term, the 20 corresponding x_{ij} for $i \neq j = 1, \ldots, 5$).

The precise way in which 2-complexes are generated as Feynman diagrams of the GFT will be illustrated in the following section. We have shown that any (local in the sense of section 3.2) spin foam model can be given a discretization-independent formulation in terms of a uniquely determined GFT theory.

9. The Barrett–Crane model and its dual GFT-formulation

In addition to providing a discretization-independent formulation of the Barrett–Crane model, the GFT formulation provides a natural completion of the definition of the model. In section 7 (equation (90)), we noted that the model on a fixed discretization is defined up to lower-dimensional simplex amplitudes such as that for faces f (dual to triangles) and edges e (dual to tetrahedra). The GFT formulation presented here resolves this ambiguity in a natural way. This normalization of the Barrett–Crane model was also obtained in [116] using similar techniques but on a fixed triangulation.

9.1. The general GFT

In this section, we introduce the general GFT action that can be specialized to define the various spin foam models described in the rest of the topical review.

9.1.1. The field theory action and its regularization. Consider the Lie group G corresponding to either $Spin(4)$ or $SL(2, \mathbb{C})$—for the GFT dual to the Riemannian or Lorentzian Barrett–Crane model, respectively. The field $\phi(x_1, x_2, x_3, x_4)$ is denoted by $\phi(x_i)$ where $i = 1, \ldots, 4$, and $x_i \in G$, symmetric under permutations of its arguments, i.e., $\phi(x_1, x_2, x_3, x_4) = \phi(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}, x_{\sigma(4)})$, for any σ permutation of four elements. Define the projectors **P** and **R** as

$$\mathbf{P}\phi(x_i) \equiv \int_G dg\,\phi(x_i g), \tag{113}$$

and

$$\mathbf{R}\phi(x_i) \equiv \int_{U^4} du_i\,\phi(x_i u_i), \tag{114}$$

where $U \subset G$ is a fixed subgroup, and dg and du are the corresponding invariant measures. The projector **P** imposes the translation invariance property (108).

Different choices of the subgroup yield different interesting GFTs. When $U = \{1\}$ ($\mathbf{R} = 1$) the GFT is dual to BF theory and we get the model of [87]. The GFT is dual to the Riemannian BC model for $U = SU(2) \subset Spin(4)$. Similarly for the GFT dual to the Lorentzian models the subgroups $U \subset SL(2, \mathbb{C})$ are $U = SU(2)$ (leaving invariant a timelike direction) or $U = SU(1, 1) \times Z_2$ (for an invariant spacelike direction); they result in two different models.

The GFT action is of the general form (104) and is simply given by

$$S[\phi] = \int_{G^4} \mathrm{d}x_i [\mathbf{P}\phi(x_i)]^2 + \frac{\lambda}{5!} \int_{G^{10}} \mathrm{d}x_{ij} \prod_{i=1}^{5} \mathbf{PRP}\phi(x_{ij}), \tag{115}$$

where $i, j = 1, \ldots, 5, i \neq j$ and $\mathrm{d}x$ is an invariant measure in G (the normalized Haar measure in G is compact).

Strictly speaking the operators $\mathbf{P}$ and $\mathbf{R}$ are projectors only when the corresponding groups are compact. Formally we have

$$\mathbf{P}^2 = vol_G \times \mathbf{P}, \qquad \text{and} \qquad \mathbf{R}^2 = vol_U \times \mathbf{R}, \tag{116}$$

where vol_G and vol_U denote the volumes of G and U, respectively. These volume factors can be taken to be 1 when the G and/or U are compact by using Haar measures. When G and/or U are non-compact the factors are infinite. This is a rather simple technical problem with which we shall deal later. Essentially one must drop redundant projectors in the functional integral.

The partition function can be computed as a perturbative expansion in Feynman diagrams $\mathcal{J}_N$, namely

$$\mathcal{Z} = \int \mathcal{D}\phi\, \mathrm{e}^{-S[\phi]} = \sum_{\mathcal{J}_N} \frac{\lambda^N}{\mathrm{sym}(\mathcal{J}_N)} A(\mathcal{J}_N), \tag{117}$$

where N is the number of vertices in $\mathcal{J}_N$ and $\mathrm{sym}(\mathcal{J}_N)$ is the standard symmetry factor.

In what follows we explain the structure of the Feynman diagrams of the theory. If we use the notation $\phi(x_{1j}) = \phi(x_{12}, x_{13}, x_{14}, x_{15})$ we can write the action as in (110) and (111)

$$\begin{aligned} S[\phi] = \frac{1}{2} \int \mathrm{d}x_i\, \mathrm{d}x_i' \phi(x_i) \mathcal{K}(x_i, x_i') \phi(x_i') \\ + \frac{\lambda}{5!} \int \mathrm{d}x_{ij} \mathcal{V}(x_{ij}) \phi(x_{1j}) \phi(x_{2j}) \phi(x_{3j}) \phi(x_{4j}) \phi(x_{5j}), \end{aligned} \tag{118}$$

where $i \neq j$ and the kinetic $\mathcal{K}(x_i, x_i')$ and interaction $\mathcal{V}(x_{ij})$ operators are explicitly given below. The kinetic operator $\mathcal{K}(x_i, x_i')$ is given by

$$\begin{aligned} \mathcal{K}(x_i, x_i') &= \sum_\sigma \int_{G^2} \mathrm{d}g'\, \mathrm{d}g'' \prod_{i=1}^{4} \delta\big(x_i g'' g'^{-1} x'^{-1}_{\sigma(i)}\big) \\ &= \sum_\sigma \int_{G^2} \mathrm{d}g'\, \mathrm{d}g \prod_{i=1}^{4} \delta\big(x_i g x'^{-1}_{\sigma(i)}\big) \\ &= vol_G \times \sum_\sigma \int_{G} \mathrm{d}g \prod_{i=1}^{4} \delta\big(x_i g x'^{-1}_{\sigma(i)}\big), \end{aligned} \tag{119}$$

where the g'' and g' integrations correspond to the action of the projectors $\mathbf{P}$ in (115). Redefining the integration variables $g = g'' g'^{-1}$ we obtain the second line in the previous equation. The vol_G factor comes from the g' integration as in (116). We regularize the kinetic

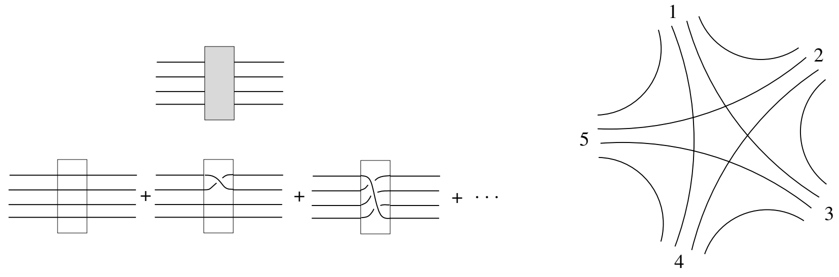


Figure 7. The structure of the propagator (right) and the interaction vertex (left). Each line represents a delta function such as those in the integrand of (120) in the case of the propagator and (123) in the case of the vertex. The shaded box in the propagator represents the sum over permutations σ in (120). This sum is over diagrams including crossings, we have represented three of these terms in the diagram below the propagator.

operator by simply dropping one of the G integrations in the previous expression, namely

$$\mathcal{K}(x_i, x_i') \equiv \sum_\sigma \int_G \mathrm{d}g \prod_{i=1}^{4} \delta\big(x_i g x'^{-1}_{\sigma(i)}\big). \tag{120}$$

P acts by projecting the field ϕ onto its 'translation-invariant' part $\mathbf{P}\phi(x_i)$. The action (115) depends only on the gauge-invariant part of the field, namely $S[\phi] = S[\mathbf{P}\phi]$ (recall (108)). The inverse of $\mathcal{K}$ (in the subspace of right-invariant fields) corresponds to itself; therefore the propagator of the theory is simply

$$\mathcal{P}(x_i, x_i') = \mathcal{K}(x_i, x_i'). \tag{121}$$

The propagator is defined by four delta functions (plus the symmetrization and the integration over the group) and it can be represented as shown in the right diagram of figure 7. The potential term (118) can be written as

$$\mathcal{V}(x_{ij}) = \frac{1}{5!} \int \mathrm{d}g_i\, \mathrm{d}g_i'\, \mathrm{d}u_{ij} \prod_{i<j} \delta\big(x_{ji}^{-1} g_i' u_{ij} g_i^{-1} g_j u_{ji} g_j'^{-1} x_{ij}\big), \tag{122}$$

where the $u_{ij} \in U$ correspond to the action of **R** in (115) and $g_i, g_i' \in G$ to the action of the corresponding **P**, respectively. It is easy to check that in the evaluation of a closed Feynman diagram the g_i' can be absorbed by redefinition of the g_i in the corresponding adjacent propagators. In this process, the g_i' drop out of the integrand and each integral over g_i' gives a vol_G factor. This means that the second **P** projector in (115) is redundant in the computation of (117). The regularization is analogous to that implemented in equation (120): we drop redundant **P**. The regularized vertex amplitude (for a vertex in the bulk of a diagram) is then defined as

$$\mathcal{V}(x_{ij}) = \frac{1}{5!} \int \mathrm{d}g_i\, \mathrm{d}u_{ij} \prod_{i<j} \delta\big(x_{ji}^{-1} u_{ij} g_i^{-1} g_j u_{ji} x_{ij}\big). \tag{123}$$

In the case of open diagrams g' remain at external legs.

There is still a redundant g_i integration in (123) which introduces a potentially infinite vol_G factor in the vertex amplitude. Note that (123) depends on the 'translational-invariant' combinations $g_i^{-1} g_j$ so that one of the g integrations is redundant. In other words, one can absorb one of the g_i by redefining the remaining four using the invariance of the measure. In the non-compact case this would yield another vol_G infinite factor. The regularization now consists of removing an arbitrary g_i from the expression of the vertex amplitude. This results in the regularization scheme proposed by Barrett and Crane in [10]. The regularization presented here can be applied to any non-compact group model on a lattice and can be regarded as a gauge-fixing condition for the internal gauge invariance (recall section 6.1). For notational simplicity we do not implement the regularization explicitly in (123). The structure of the vertex is represented in the left diagram in figure 7. Each of the lines represents a delta function appearing in the integrand of (123) (compare with (112)).

The Feynman diagrams of the theory are obtained by connecting the five-valent vertices with propagators (see figure 7). At the open ends of propagators and vertices there are the four group variables x corresponding to the arguments of the field (in addition to the integration variables g_i and u_{ij} coming from $\mathbf{P}$ and $\mathbf{R}$, respectively). For a fixed permutation σ in each propagator, one can follow the sequence of delta functions with common arguments across vertices and propagators. On a closed graph, each such sequence must close. By associating a surface with each such sequence of propagators, we construct a 2-complex $\mathcal{J}$ [11]. Thus, by expanding in Feynman diagrams and in the sum over permutations in (119), we obtain a sum over 2-complexes. Each 2-complex is given by a certain vertex–propagator topology plus a fixed choice of a permutation on each propagator.

9.1.2. Evaluation in configuration space. Combining (121) and (123) and integrating over internal variables x_{ij} the sequence of delta functions associated with a face $f \in \mathcal{J}$ reduces to a single delta function. Denoting by $v_1 \ldots v_n$ the ordered set of n vertices bounding f, and e_{ij} the edge connecting v_i with v_j, the delta function corresponding to the face f becomes

$$\delta\big(\big[u^f_{v_{1n}} g^{-1}_{v_{1n}} g_{v_{12}} u^f_{v_{12}}\big] g_{e_{12}} \big[u^f_{v_{21}} g^{-1}_{v_{21}} \hat{g}_{v_{23}} u^f_{v_{23}}\big] g_{e_{23}} \cdots g_{e_{(n-1)n}} \big[\hat{u}^f_{v_{n(n-1)}} g^{-1}_{v_{n(n-1)}} g_{v_{n1}} u^f_{v_{n1}}\big] g_{e_{n1}}\big), \tag{124}$$

where $g_{e_{ij}} \in G$ is the integration variable in (120) associated with the edge e_{ij}, $g_{v_{ij}} \in G$ denotes the integration variable in (123) associated with the leg ij of the vertex v_i connected to the edge e_{ij}, and $u^f_{v_{ij}} \in U$ is the corresponding subgroup integration in one of the delta functions in (123). We use the superscript f to emphasize that the u appear on single delta functions in (123) and therefore do not contribute to more than one face. In the previous equation, the product of group elements between brackets corresponds to the vertex contribution to the face (see (123)) while the $g_{e_{ij}}$ comes from the corresponding propagators (121), also we have that $g_{e_{ij}} = g^{-1}_{e_{ji}}$. Using (124) the amplitude $A(\mathcal{J})$ of an arbitrary closed 2-complex becomes a multiple integral of the form

$$A(\mathcal{J}) = \int \prod_e \mathrm{d}g_e \prod_v \mathrm{d}g_v\, \mathrm{d}g_v \prod_f \mathrm{d}u^f_v\, \mathrm{d}u^f_v\, \delta\left(\big[u^f_{v_{1n}} g^{-1}_{v_{1n}} g_{v_{12}} u^f_{v_{12}}\big]\, g_{e_{12}}\right.$$
$$\left. \times \big[u^f_{v_{21}} g^{-1}_{v_{21}} \hat{g}_{v_{23}} u^f_{v_{23}}\big]\, g_{e_{23}} \cdots g_{e_{(n-1)n}} \big[\hat{u}^f_{v_{n(n-1)}} g^{-1}_{v_{n(n-1)}} g_{v_{n1}} u^f_{v_{n1}}\big] g_{e_{n1}}\right). \tag{125}$$

Note that if the subgroup $U = \{1\}$ the previous expression coincides with the BF amplitude (61) if we make the change of variables $g'_{e_{ij}} = g_{v_{ij}} g_{e_{ij}} g^{-1}_{v_{ji}}$. For $U = \{1\}$ we recover Ooguri's GFT [87].

9.1.3. The spin foam representation. The spin foam representation for the amplitude can be obtained by expanding the delta functions in terms of irreducible unitary representations of G, namely

$$\delta(g) = \sum_\rho \Delta_\rho \mathrm{Tr}[\rho(g)], \tag{126}$$

where ρ labels unitary irreducible representations and the rest of the notation is that of (31) when G is compact. In the non-compact case representations are infinite-dimensional so a formally equivalent expression holds where Δ_ρ corresponds to the so-called Pancharel measure [110].

Using (126) in (124) the face contribution is

$$\sum_\rho \Delta_\rho \mathrm{Tr}\left[\rho\left(u^f_{v_{1n}}\right)\cdot\rho\left(g^{-1}_{v_{1n}}g_{v_{12}}\right)\cdot\rho\left(u^f_{v_{12}}\right)\cdot\rho\left(g_{e_{12}}\right)\cdot\rho\left(u^f_{v_{21}}\right)\cdot\rho\left(g^{-1}_{v_{21}}g_{v_{23}}\right)\right.$$
$$\left.\cdot\,\rho\left(u^f_{v_{23}}\right)\cdot\rho\left(g_{e_{23}}\right)\cdots\rho\left(u^f_{v_{n(n-1)}}\right)\cdot\rho\left(g^{-1}_{v_{n(n-1)}}g_{v_{n1}}\right)\cdot\rho\left(u^f_{v_{n1}}\right)\cdot\rho\left(g_{e_{n1}}\right)\right], \tag{127}$$

where we have dropped the integration symbols for simplicity. The u appear just once per face so we can perform the u integrations independently of other faces. Note that $\int \mathrm{d}u\rho(u)$ is the projection onto the invariant subspace under the action of U in the Hilbert space $\mathcal{H}_\rho$. When $U = \{1\}$ the projector is the identity. In the other cases the subspace turns out to be one dimensional. Consequently, the projector can be written as

$$\int_U \mathrm{d}u\rho(u) = |w_\rho\rangle\langle w_\rho|, \tag{128}$$

where $|w_\rho\rangle$ is the corresponding invariant vector. Equation (127) becomes

$$\sum_\rho \Delta_\rho \left\langle w_\rho\right|\rho\left(g^{-1}_{v_{1n}}g_{v_{12}}\right)w_\rho\rangle\langle w_\rho|\rho\left(g_{e_{12}}\right)w_\rho\rangle\langle w_\rho|\rho\left(g^{-1}_{v_{21}}g_{v_{23}}\right)w_\rho\rangle\cdots\langle w_\rho|\rho\left(g_{e_{n1}}\right)w_\rho\rangle. \tag{129}$$

There are two types of terms in the previous equation: those corresponding to edges for which the representation ρ is evaluated at the single $g_{e_{ij}}$ group variable, and those corresponding to vertices for which ρ is evaluated at the product $g^{-1}_{v_{ij}}g_{v_{ik}}$. Let us integrate over the edge variables first. The group element g_e associated with an edge $e \in \mathcal{J}$ appears four times as there are four delta functions in the propagator (121). Therefore, integrals over such g have the general form

$$A_e\left(\rho^1_e,\ldots,\rho^4_e\right) = \int_G \langle w_{\rho^1}|\rho^1(g)w_{\rho^1}\rangle\langle w_{\rho^2}|\rho^2(g)w_{\rho^2}\rangle\langle w_{\rho^3}|\rho^3(g)w_{\rho^3}\rangle\langle w_{\rho^4}|\rho^4(g)w_{\rho^4}\rangle, \tag{130}$$

and define the edge amplitude. If we define the *kernel* $K_\rho(g)$ as

$$K_\rho(g) \equiv \langle w_\rho|\rho(g)w_\rho\rangle, \tag{131}$$

then using the subgroup invariance of the $|w_\rho\rangle$ the previous equation becomes

$$A_e(\rho_1,\ldots,\rho_4) = vol_U \times \int_{G/U} \mathrm{d}y K_{\rho_1}(y)K_{\rho_2}(y)K_{\rho_3}(y)K_{\rho_4}(y), \tag{132}$$

where the integration is over the homogeneous space G/U. Integration over the group elements associated with vertices are of the general form

$$A_v(\rho_{ik}) = \int_{(G/U)^4}\prod_{i=1}^5 \mathrm{d}y_i \prod_{i<k} K_{\rho_{ik}}(y_i, y_k).$$

This corresponds to the vertex amplitude. When U is non-compact the preceding expressions have to be regularized as in subsection 9.1.1 by dropping an arbitrary y integration. The amplitude of a 2-complex $\mathcal{J}$ is then given by

$$A(\mathcal{J}) = \sum_{\mathcal{C}_f:\{f\}\to\{\rho\}}\prod_{f\in\mathcal{J}}\Delta_{\rho_f}\prod_{e\in\mathcal{J}}A_e(\rho_1,\ldots,\rho_4)\prod_{v\in\mathcal{J}}A_v(\rho_1,\ldots,\rho_{10}). \tag{133}$$

As we shall see in the following, the presence of **R** in (115) can be interpreted as imposing the Barrett–Crane constraints on the GFT dual to BF theory.

9.1.4. Boundaries. The amplitude of an open diagram, that is, a diagram with a boundary, is a function of the variables on the boundary, as for conventional QFT Feynman diagrams. The boundary of a 2-complex is given by a graph. To start with, the amplitude of the open diagram is a function of four group arguments per each external leg. However, consider a surface of an open 2-complex and the link ab of the boundary graph that bounds it. Let a and b be the nodes on the boundary graph that bounds ab. The surface determines a sequence of delta functions that starts with one of the group elements in a, say g_a, and ends with one of the group elements in b, say g_b. By integrating internal variables, all these delta functions can be contracted to a single one of the form $\delta\big(g_a \cdots g_b^{-1}\big) = \delta\big(g_b^{-1} g_a \cdots\big)$. We can thus define the group variable $\rho_{ab} = g_b^{-1} g_a$, naturally associated with the link ab, and conclude that the amplitude of an open 2-complex is a function $A(\rho_{ab})$ of one group element per each link of its boundary graph. $A(\rho_{ab})$ is gauge invariant as can be easily checked, and ρ_{ab} represents the discrete boundary connection in the language of section 3.2.

In 'momentum space', boundary degrees of freedom are encoded in spin network states. That is, if s is a spin network given by a colouring of the boundary graph, then

$$A(s) = \int \mathrm{d}\rho_{ab} \bar{\psi}_s(\rho_{ab}) A(\rho_{ab}), \tag{134}$$

where $\psi_s(\rho_{ab})$ is the spin network function [26, 27]. The formula can be inverted

$$A(\rho_{ab}) = \sum_s A(s) \psi_s(\rho_{ab}), \tag{135}$$

where the sum is performed over all spin network states defined on the given boundary graph. The argument presented here applies in general for any GFT of the kind introduced in section 8. In the case of the GFTs defined here $\rho_{ab} = g_b^{-1} g_a$ will appear between U integrations. So strictly speaking, for each $\rho_{ab} = g_b^{-1} g_a$ there correspond two points y_a and y_b in the homogeneous space G/U and one cannot say that the amplitude depends on the G connection.

9.2. GFT dual to the Riemannian Barret–Crane model

The representation theory of $Spin(4)$ is particularly simple due to the fact that $Spin(4) = SU(2) \times SU(2)$. Unitary irreducible representations of ρ of $Spin(4)$ are labelled by two half-integers j_l and j_r, and are given by the tensor product of unitary irreducible representations of $SU(2)$, namely

$$\rho_{j_l j_r} = j_l \otimes j_r. \tag{136}$$

In terms of representation matrices we have

$$R^{j_l j_r}(g)_{mm'qq'} = D^{j_l}_{mm'}(g_r) D^{j_r}_{qq'}(g_l), \tag{137}$$

where $g = (g_l, g_r) \in Spin(4)$, and $D^{j_l}_{mm'}$ are $SU(2)$ representation matrices. The subgroup U of previous section is taken as $U = SU(2) \subset Spin(4)$ defined by the diagonal action on $Spin(4)$ as follows:

$$gu \equiv (g_l u, g_r u) \qquad \text{for} \quad u \in U. \tag{138}$$

The projector (128) can be explicitly evaluated

$$\int_U \mathrm{d}u \rho_{j_l j_r}(u) = \int \mathrm{d}u D^{j_l}_{mm'}(u) D^{j_r}_{qq'}(u) = \frac{\delta_{j_l j_r^*}}{2j_l + 1} \delta_{mq} \delta_{m'q'} = |w_{j_l}\rangle\langle w_{j_l}|, \tag{139}$$

where we have used the orthonormality of $SU(2)$ representation matrices (footnote 10). The previous equation confirms that the invariant subspace for each $\rho_{j_l j_r}$ is one dimensional as anticipated above. An immediate consequence of (139) is that the kernel K_ρ in (131) becomes

$$K_j(g) = \langle w_j|\rho_{j'j}(g)w_j\rangle = \frac{\delta_{j'j^*}}{2j+1}\mathrm{Tr}[D^j(g_l g_r)]. \tag{140}$$

The factor $\delta_{j'j}$ restricts the representation to the simple representations $j_l = j_r^* = j$.

The projector **R** in (115) imposes the Barrett–Crane simplicity constraints of equation (83)! Simple representations appear in the harmonic analysis of functions on the 3-sphere, S_3. The fact that we encounter them here is not surprising since S_3 is the homogeneous space $S_3 = Spin(4)/SU(2)$ under the $SU(2)$ diagonal insertion (138). The presence of R in (115) projects out those modes that are not simple or *spherical*. Indeed, $\langle w_j|\rho_{j'j^*}(g)w_j\rangle$ can be thought of as a function on S_3: $\langle w_j|\rho_{j'j^*}(g)w_j\rangle$ depends on the product $g_l g_r \in SU(2)$ which is isomorphic to S_3. We will see below that in fact **R** imposes all the constraints of section 7.3 fixing in addition the value of the so far undetermined edge amplitude A_e.

The invariant measure on $SU(2)$ can be written as a measure on S_3 induced by the isomorphism $SU(2) \to S_3$. We can parametrize $h \in SU(2)$ as $h = y^\mu_{(h)}\tau_\mu$ where $\tau_k = \mathrm{i}\sigma_k$ for $k = 1, 2, 3$, $\{\sigma_k\}$ is the set of Pauli matrices and $\tau_0 = 1$. $h \in SU(2)$ implies $y^\mu_{(h)} y_{(h)\mu} = 1$ (indices are lowered and raised with δ^μ_ν), i.e., $y \in S_3$. In this parametrization the $SU(2)$ Haar measure becomes

$$\mathrm{d}h \to \mathrm{d}y = \frac{1}{\pi^2}\,\mathrm{d}y^4\delta(y^\mu y_\mu - 1). \tag{141}$$

We can simplify the measure using spherical coordinates for which an arbitrary point $y \in S^3$ is written

$$y = (\cos(\psi), \sin(\psi)\sin(\theta)\cos(\phi), \sin(\psi)\sin(\theta)\sin(\phi), \sin(\psi)\cos(\theta)), \tag{142}$$

where $0 \leqslant \psi \leqslant \pi$, $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$. The Haar measure becomes

$$\mathrm{d}y = \frac{1}{2\pi^2}\sin^2(\psi)\sin(\theta)\,\mathrm{d}\psi\,\mathrm{d}\theta\,\mathrm{d}\phi. \tag{143}$$

Using the known formula for the trace of $SU(2)$ representation matrices it is easy to check that (140) becomes

$$K_j(\psi_y) = \frac{\delta_{j'j}}{2j+1}\frac{\sin((2j+1)\psi_y)}{\sin(\psi_y)}, \tag{144}$$

where ψ_h is the value of the coordinate ψ in (142) corresponding to $y \in SU(2)$. Note that we have rediscovered the kernel (93) which together with equation (94) establishes the claimed duality between the Barrett–Crane model and the GFT of this section.

Now, the general edge and vertex amplitudes ((132) and (133)) reduce to multiple integrals over S_3. These can be interpreted as the evaluation of diagrams on the sphere with the propagator given by (144). This evaluation is referred to as *relativistic spin network evaluation* in the literature. In the case of the edge amplitude the evaluation corresponds to a θ_4 graph with four links labelled by the four spins $j_1, \ldots, j_4$. In the case of the vertex amplitude the evaluation corresponds to the γ_5 spin network of figure 4 for which the ten links are labelled by the corresponding ten spins $j_1, \ldots, j_{10}$. There is no intertwiner label at nodes as the projection **R** fixes the intertwiners to the unique Barrett–Crane one.

The amplitude $A_e(j_1, \ldots, j_4)$ can be computed using (140) and (33), it follows that

$$A_e(j_1, \ldots, j_4) = \frac{\dim[\mathrm{Inv}(j_1 \otimes j_2 \otimes j_3 \otimes j_4)]}{\dim[j_1 \otimes j_2 \otimes j_3 \otimes j_4]}, \tag{145}$$

where dim [Inv($j_1 \otimes j_2 \otimes j_3 \otimes j_4$)] denotes the dimension of the trivial component of $\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} \otimes \mathcal{H}_{j_3} \otimes \mathcal{H}_{j_4}$ and dim [$j_1 \otimes j_2 \otimes j_3 \otimes j_4$] the dimension of the full space[30]. So finally,

$$A(\mathcal{J}) = \sum_{\mathcal{C}_f:\{f\}\to\{j\}} \prod_{f\in\mathcal{J}} (2j_f+1)^2 \prod_{e\in\mathcal{J}} A_e(j_1,\ldots,j_4) \prod_{v\in\mathcal{J}} A_v(j_1,\ldots,j_{10}), \tag{146}$$

where the vertex amplitude is given by (133) using (144). This is the Barrett–Crane model (90) where the values of A_e and A_f have been fixed by the GFT formulation.

9.2.1. Finiteness. In this section, we prove that the infinite spin sum in (146) is convergent if the 2-complex $\mathcal{J}$ is finite and non-degenerate. In order to do this we need to construct bounds for the terms in the sum (edge and vertex amplitudes). These bounds are provided by the following lemmas.

Lemma 9.1 [18]**.** *For any subset of κ elements $j_1,\ldots,j_\kappa$ out of the corresponding four representations appearing in $A_e(j_1,\ldots,j_4)$, the following bounds hold:*

$$|A_e(j_1,\ldots,j_4)| \leqslant \frac{1}{\prod_{i=1}^{\kappa}(2j_i+1)^{\alpha_\kappa}}, \qquad \textit{where} \quad \alpha_\kappa = \begin{cases} 1 & \textit{for} \quad \kappa \leqslant 3 \\ \frac{3}{4} & \textit{for} \quad k = 4 \end{cases}.$$

The inequality for $\kappa \leqslant 3$ is sharp.

The proof of this lemma is elementary and can be found in [18].

Lemma 9.2 [18]**.** *For any four spins $j_1,\ldots,j_4$ labelling links converging at the same node in the relativistic spin network corresponding to $A_v(j_1,\ldots,j_{10})$ the following bounds hold:*

$$|A_v(j_1,\ldots,j_{10})| \leqslant \frac{1}{(2j_1+1)(2j_2+1)(2j_3+1)(2j_4+1)},$$

from which follows

$$|A_v(j_1,\ldots,j_{10})| \leqslant \frac{1}{((2j_1+1)\cdots(2j_{10}+1))^{2/5}}.$$

Definition 9.1. *A 2-complex $\mathcal{J}$ is said to be degenerate if it contains some faces bounded by only one or two edges.*

Theorem 9.1 [18]**.** *The state sum $A(\mathcal{J})$, (146), converges for any non-degenerate 2-complex $\mathcal{J}$.*

Proof. The amplitude (146) can be bounded in the following way:

$$|A(\mathcal{J})| \leqslant \prod_{f\in\mathcal{J}} \sum_{j_f} ((2j_f+1))^{2-\frac{3}{4}n_f-\frac{4}{10}n_f} = \prod_{f\in\mathcal{J}} \sum_{j_f} ((2j_f+1))^{2-\frac{23}{20}n_f}, \tag{150}$$

where n_f denotes the number of edges of the corresponding face, and we have used the fact that the number of edges equals the number of vertices in a face of $\mathcal{J}$. The term $(2j+1)^2$ in (150) comes from the face amplitude, $(2j+1)^{-\frac{3}{4}n_f}$ from lemma 9.1 ($\kappa = 4$), and $(2j+1)^{-\frac{4}{10}n_f}$ from lemma 9.2 ($\kappa = 10$). Note that if the 2-complex contains faces with more than two edges the previous bound for the amplitude is finite, since the sum on the RHS of the previous equation converges for $n_f > 2$. □

[30] There was a mistake in the computation of the edge amplitude in [12] due to the propagation of a typo in [11]. The erroneous expression of the edge amplitude contained dim [$j_1 \otimes j_2 \otimes j_3 \otimes j_4$]^2 in the denominator.

The fact that the sum over spins converges for a fixed (non-degenerate) 2-complex is a very encouraging result. It means that for a fixed 2-complex the sum over the infinite set of 4-geometry configurations, represented by the corresponding spin foams, is well defined. Even though the result was obtained in the context of the GFT formulation, it can be applied to the model defined on a simplicial decomposition of the spacetime manifold. This is because (146) refers to the amplitude of a single Feynman diagram. In this context, the result is very encouraging and appears as a first necessary step for the construction of the discretization-independent generalized projector and physical scalar product by studying the refinement limit of section 6.2. On the other hand, if the discretization-independent formulation is to be obtained directly by means of the GFT formulation, this result means that the perturbative expansion in λ (105) is finite order by order. The previous statement has not been completely proved as degenerate 2-complexes, for which our proof does not hold, will appear in the λ expansion. However, we expect these diagrams to be finite also. A proof of this is lacking at the moment but it should be feasible by strengthening the bounds we have used in this section.

9.3. Lorentzian models

Unitary irreducible representations of $SL(2,\mathbb{C})$ appearing in the general expression (126) are those in the so-called *principal* series. They are labelled by a natural number n and a positive real number ρ. Unitary irreducible representations of $SL(2,\mathbb{C})$ are infinite dimensional. Those in the principal series are defined by their action on the linear space $\mathcal{D}_{n\rho}$ of homogeneous functions of degree $\left(\frac{n+\mathrm{i}\rho}{2}-1, \frac{-n+\mathrm{i}\rho}{2}-1\right)$ of two complex variables z_1 and z_2. Due to the homogeneity properties of the elements $\mathcal{D}_{n\rho}$ they can be characterized by giving their value on the sphere $|z_1|^2+|z_2|^2=1$. The latter is isomorphic to $SU(2)$. The so-called canonical basis is defined in terms of the functions on $SU(2)$ and is well suited for studying the following model. The relevant facts about $SL(2,\mathbb{C})$ representation theory and the notation used in this section can be found in the appendix of [110] (for all about $SL(2,\mathbb{C})$ representation theory, see [117]).

9.3.1. GFT dual to the Lorentzian Barrett–Crane model. Equation (128) becomes, in this case,

$$\int_U \mathrm{d}u\rho_{n\rho}(u) = \int_{SU(2)} \mathrm{d}u\, D^{n\rho}_{\ell m\ell' m'}(u) = \delta_{n0}\delta_{\ell 0}\delta_{\ell' 0} = |w_{n\rho}\rangle\langle w_{n\rho}|, \tag{151}$$

in terms of the canonical basis [110]. The $SU(2)$-invariant vectors $|w_{n\rho}\rangle$ are given by $\langle z_1, z_2|w_{n\rho}\rangle = \delta_{n0}(|z_1|^2+|z_2|^2)^{\frac{\mathrm{i}}{2}\rho-1}$ which is indeed a homogeneous function of degree $\left(\frac{\mathrm{i}\rho}{2}-1, \frac{\mathrm{i}\rho}{2}-1\right)$.

An immediate consequence of (151) is that the kernel K_ρ in (131) becomes

$$K_{n\rho}(g) = \langle w_{n\rho}|\rho_{n\rho}(g)w_{n\rho}\rangle = \delta_{n0}D^{n\rho}_{0000}(u). \tag{152}$$

The kernel $\langle w_{0\rho}|\rho_{0\rho}(g)w_{0\rho}\rangle$ can be thought of as a function on $H^+ = SL(2,\mathbb{C})/SU(2)$ (the upper sheet of a two-sheeted hyperboloid in Minkowski space). An arbitrary point $y \in H^+$ can be written in hyperbolic coordinates as

$$y = (\cosh(\eta), \sinh(\eta)\sin(\theta)\cos(\phi), \sinh(\eta)\sin(\theta)\sin(\phi), \sinh(\eta)\cos(\theta)), \tag{153}$$

where $0 \leqslant \eta \leqslant \infty$, $0 \leqslant \theta \leqslant \pi$, and $0 \leqslant \phi \leqslant 2\pi$. The invariant measure in these coordinates is

$$\mathrm{d}y = \frac{1}{2\pi^2}\sinh^2(\eta)\sin(\theta)\,\mathrm{d}\eta\,\mathrm{d}\theta\,\mathrm{d}\phi. \tag{154}$$

In these coordinates, (152) becomes

$$K_\rho(\eta_y) = \frac{2}{\rho}\frac{\sin\left(\frac{1}{2}\rho\eta_y\right)}{\sinh(\eta_y)}, \tag{155}$$

where η_y is the value of the coordinate η corresponding to $y \in H^+$ (this the generalization of (93) proposed in [10]). Finally, the amplitude of an arbitrary diagram (133) becomes

$$A(J) = \int_{\mathcal{C}_f:\{f\}\to\{\rho\}} \prod_{f\in J} \rho_f^2 \, \mathrm{d}\rho_f \prod_{e\in J} A_e(\{\rho_e\}) \prod_{v\in J} A_v(\{\rho_v\}), \tag{156}$$

where the formal sum in the general expression (133) becomes a multiple integral over the colouring $\mathcal{C}_f : \{f\} \to \{\rho\}$ of faces, and the weight Δ_ρ now is given by the Pancherel measure, $\rho^2\,\mathrm{d}\rho$, of $SL(2,\mathbb{C})$. The vertex amplitude was proposed in [10], the full state sum amplitude including the previous normalization was obtained in [13].

9.3.2. Finiteness. The amplitude of a non-degenerate 2-complex turns out to be finite as in the Euclidean case. We state and prove the main result. Some of the following lemmas are stated without proof; the reader interested in the details is referred to the references. As in the Riemannian case the main idea is to construct appropriate bounds for the fundamental amplitudes.

Lemma 9.3 [19, 20]. *For any subset of κ elements $\rho_1, \ldots, \rho_\kappa$ out of the corresponding four representations appearing in $A_e(\rho_1, \ldots, \rho_4)$, the following bounds hold:*

$$|A_e(\rho_1, \ldots, \rho_4)| \leqslant \frac{C_\kappa}{\left(\prod_{i=1}^{\kappa} \rho_i\right)^{\alpha_\kappa}}, \qquad \textit{where} \quad \alpha_\kappa = \begin{cases} 1 & \textit{for} \quad \kappa \leqslant 3 \\ \frac{3}{4} & \textit{for} \quad k = 4 \end{cases},$$

for some positive constant C_κ.

Lemma 9.4 [118]. *$A_e(\rho_1, \ldots, \rho_4)$ and $A_v(\rho_1, \ldots, \rho_{10})$ are bounded by a constant independent of the ρ.*

Lemma 9.5 [19, 20]. *For any subset of κ elements $\rho_1, \ldots, \rho_\kappa$ out of the corresponding ten representations appearing in $A_v(\rho_1, \ldots, \rho_{10})$ the following bounds hold:*

$$|A_v(\rho_1, \ldots, \rho_{10})| \leqslant \frac{K_\kappa}{\left(\prod_{i=1}^{\kappa} \rho_i\right)^{\frac{3}{10}}}$$

for some positive constant K_κ.

Theorem 9.2 [19, 20]. *The state sum $A(\mathcal{J})$, (156), converges for any non-degenerate 2-complex $\mathcal{J}$.*

Proof. We divide each integration region $\mathbb{R}^+$ into the intervals $[0, 1)$, and $[1, \infty)$ so that the multiple integral decomposes into a finite sum of integrations of the following types.

i. All the integrations are in the range $[0, 1)$. We denote this term $T(F, 0)$, where F is the number of 2-simplexes in the triangulation. This term in the sum is finite by lemma 9.4
ii. All the integrations are in the range $[1, \infty)$. This term $T(0, F)$ is also finite since, using lemmas 9.3 and 9.5 for $\kappa = 4$, and $\kappa = 10$ respectively, we have

$$T(0, F) \leqslant \prod_f \int_{\rho_f=1}^{\infty} \mathrm{d}\rho_f \, \rho_f^{2-\frac{3}{4}n_e-\frac{3}{10}n_v} \leqslant \left(\int_{\rho_f=1}^{\infty} \mathrm{d}\rho_f \, \rho_f^{-\frac{46}{40}}\right)^F < \infty.$$

iii. m integrations in $[0, 1)$, and $F - m$ in $[1, \infty)$. In this case $T(m, F - m)$ can be bounded using lemmas 9.3 and 9.5 as before. The idea is to choose the appropriate subset of representations in the bounds (and the corresponding values of κ) so that only the $m - F$ representations integrated over $[1, \infty)$ appear in the corresponding denominators. Since this is clearly possible, the $T(m, F)$ terms are all finite.

We have bounded $\mathcal{Z}$ by a finite sum of finite terms which concludes the proof. □

The 'scaling' behaviour of amplitudes in the model is very similar to that of the Riemannian version of the previous section. If one makes the substitution $(2j+1) = -\mathrm{i}\rho$ and $\psi_y = \mathrm{i}\eta_y$ in the Riemannian expression for the kernel (144) one obtains the Lorentzian expression (155). In addition, with the same substitution the face amplitude $(2j+1)^2$ in (146) takes the form ρ^2 of the face amplitude in (156). From this viewpoint the finiteness of the Lorentzian model is intimately related to the finiteness of the Riemannian counterpart.

The non-compactness of $SL(2,\mathbb{C})$ introduces potential divergences due to the fact that $vol_{SL(2,\mathbb{C})} = \infty$. This problem can be dealt with in the GFT formulation by using the appropriate gauge-fixing conditions as explained in section 9. If we define the model on a simplicial decomposition the same gauge fixing can be implemented.

9.3.3. A new Lorentzian model. In the Euclidean case there was only one way of selecting the subgroup U of group elements leaving invariant a fixed direction in Euclidean spacetime. In the Lorentzian case there are two possibilities. The case in which that direction is spacelike was treated in the previous section. When the direction is timelike the relevant subgroup is $U = SU(1,1) \times \mathrm{Z}_2$.

This case is more complicated due to the non-compactness of U. Consequently, one has to deal with additional infinite volume factors of the form U-volume in (116). Another consequence is that the invariant vectors defined in (128) are now distributional and therefore no longer normalizable. All these make more difficult the convergence analysis performed in the previous models and the issue of finiteness remains open.

On the other hand, the model is very attractive as its state sum representation contains simple representations in both the continuous and discrete series. As pointed out in [10] and discussed in the following section one would expect both types of simple representations to appear in a model of Lorentzian quantum gravity.

The difficulties introduced by the non-compactness of U make the calculation of the relevant kernels (131) more involved. No explicit formulae are known and they are defined by integral expressions. We will not derive these expressions here. A complete derivation can be found in [110]. The idea is to use harmonic analysis on the homogeneous space $SL(2,\mathbb{C})/SU(1,1) \times \mathrm{Z}_2$ which can be realized as the one-sheeted hyperboloid $y^\mu y^\nu \eta_{\mu\nu} = -1$, where y and $-y$ are identified (imaginary Lobachevskian space, from now on denoted by H^-). The kernels correspond to eigenfunctions of the massless wave equation on that space.

The corresponding kernel is given by the following expression:

$$K_{\rho,n}(x,y) = \int_{C^+} \mathrm{d}\omega \Big(\delta_{n,0}\, |y^\nu \xi_\nu|^{\mathrm{i}\frac{\rho}{2}-1}\, |x^\nu \xi_\nu|^{-\mathrm{i}\frac{\rho}{2}-1} + \delta(\rho)\delta_{n,4k} \frac{32\pi\, \mathrm{e}^{-2\mathrm{i}k[\Theta(x,y)]}}{k} \delta(x^\nu \xi_\nu)\delta(y^\nu \xi_\nu)\Big), \tag{159}$$

where $x, y \in H^-$ and $\xi \in C^+$ is a normalized future-pointing null vector in Minkowski spacetime. The integration is performed on the unit sphere defined by these vectors with the standard invariant measure $\mathrm{d}\omega$.

As in the previous cases the expressions for the edge and vertex amplitudes (equations (132) and (133)) reduce to integrals on the hyperboloid H^-. The expression for the amplitude (133) of an arbitrary diagram is [110]

$$A(J) = \sum_{n_f} \int_{\rho_f} \mathrm{d}\rho_f \prod_f \left(\rho_f^2 + n_f^2\right) \prod_e A_e(\{n_e\}, \{\rho_e\}) \prod_v A_v(\{n_v\}, \{\rho_v\}),$$

where now there is a summation over the discrete representations n and a multiple integral over the continuous representations ρ. For the discrete representations the triangle area spectrum takes the form (97) if we define $k = j + 1/2$. The finiteness properties of this model have not been studied so far. There is a relative minus sign between the continuous and discrete eigenvalues of the area squared operators that has been interpreted as providing a notion of microcausality in [110]. It would be interesting to study this in connection with the models of section 5.5.

10. Discussion

We conclude this topical review with some remarks and a discussion of recent results and future perspectives in the subject. As we have devoted the second part of this topical review almost exclusively to the Barrett–Crane model we start by discussing open issues in this context. We go to more general open questions towards the end of this section.

1. *Normalization*. The Barrett–Crane model is certainly the most well-studied spin foam model for four-dimensional quantum gravity. However, as we pointed out at the end of section 7, the definition of the measure is incomplete since there remains to determine the values of lower-dimensional simplex amplitudes such as the edge (tetrahedron) amplitude A_e. Although, the normalization (145)–(146) of the Barrett–Crane model can be fixed in a natural way using the GFT formulation, or using similar (subgroup averaging) arguments [116, 119, 120], one would really like to understand this prescription in relation to the continuum theory.

   In section 7, we have shown that the Barrett–Crane model can in fact be interpreted as a 'lattice' formulation of the gravity sector of the *SO*(4) Plebanski theory; this conclusion can be extended to the Lorentzian theory without further complications. This close relationship between the model and a continuum action suggested the possibility of determining the correct spin foam measure by comparison with the formal continuum path phase space path integral measure of Plebanski's theory. If the value of the edge amplitude (145) could be found in this way this would provide further evidence for the GFT–gravity duality and strengthen the finiteness results. These considerations apply also for other spin foam models, in particular, to those defined by constraining BF theory. A general prescription for the construction of an anomaly-free measure for spin foam models is investigated in [85].
2. *Numerical results*. A very attractive property of the normalization (146) is, however, the finiteness of transition amplitudes on a fixed discretization. Having a well-defined model—for both Riemannian and Lorentzian gravity—has opened for the first time the possibility for numerical explorations. An important step towards the numerical evaluation of the BC model is the development of efficient algorithms for the computation of the $10j$-symbols [121]. Numerical calculations by Baez *et al* [93] show that the sum over spin foams in our normalization of the Barrett–Crane model converges very fast so that amplitudes are dominated by spin foams where most of the faces are labelled by zero spin. The leading contribution comes from spin foams made up from isolated 'bubbles of

geometry'. This problem is, however, not present in the Lorentzian version (156) where simple representations are in the continuum part of the spectrum. Baez, Christensen, Halford and Tsang propose modifications of the normalization which still yield finite amplitudes but avoid this puzzling feature. However, these modifications consist of *ad hoc* definitions of the values of edge and face amplitudes. A clear understanding of this would have to come with the understanding of the normalization issues discussed above.

If we take seriously the normalization (145), the previous result might not be so worrisome after all. According to section 7.6, the asymptotic expression of the vertex amplitude shows that for large spins it is dominated by degenerate 4-simplexes. These configurations are, however, strongly suppressed by the form of the edge amplitude (145) in our normalization. Assuming that for small spin values the amplitudes behave appropriately, large-spin behaviour might not be physically relevant after all.

As pointed out by Smolin [122], one could neglect spin-zero representations according to the argument that they represent 'zero geometry states' (all geometric operators vanish on the spin-zero subspace). If one does so, then transition amplitudes are dominated by the spin-one-half representations and the model becomes very similar to a model of dynamical triangulations: in dynamical triangulations one sums over triangulations whose 1-simplexes have the same definite length, while here the amplitudes are dominated by triangulations where triangles have the same fixed area given by (97) for $j = 1/2$. However, one should also point out an important difference: in our case, configurations are weighted with complex amplitudes $\exp(\mathrm{i}S)$ which have no (*a priori*) connection to the Euclidean amplitude $\exp(-S)$ of dynamical triangulations. In the Lorentzian models the analogy simply disappears as representation labels are continuous.

3. *The connection with the canonical picture*. We began this topical review motivating the spin foam approach from the perspective of loop quantum gravity. Can we establish a rigorous connection between the spin foam models presented here and the canonical formulation? In all the models (for Riemannian gravity) introduced in section 5, with the exception of the Barrett–Crane model, this relationship is manifest at the kinematical level as the boundary states are given by $SU(2)$ spin networks. This kinematic connection exists by construction in these cases.

   What happens in the case of the Barrett–Crane model? The Riemannian BC model is defined in terms of $Spin(4)$, so naturally one could expect the boundary data to be given by $Spin(4)$ spin networks (labelled by two half-integers (j_ℓ, j_r)). The simplicity constraints impose $j_\ell = j_r = j$ which could be interpreted as the existence of an underlying $SU(2)$ connection. However, it can be shown that the boundary data of the model cannot be interpreted as given by an $SU(2)$ connection (recall the discussion of section 9.1.4). Similarly, in the Lorentzian Barrett–Crane model boundary states could be naively related to $SL(2, \mathbb{C})$ spin networks. However, similar considerations as in the Riemannian sector show that the simplicity constraints (applied in the style of Barrett–Crane) reduce the boundary states in a way that cannot be interpreted as a boundary connection. Alexandrov [123, 124] has been studying the possibility of defining an $SL(2, \mathbb{C})$ connection canonical formulation of general relativity. The second class constraints of the theory—related to the simplicity constraints of Plebanski's formulation—are dealt with using the Dirac bracket in terms of which the original connection becomes non-commutative. In this sense there is no genuine or standard spin network representation of states. It seems that such an approach could lead to the understanding of the relationship of the Barrett–Crane model and a canonical formulation of gravity. There are in fact some indications of this [125, 126].

Another difficulty in making contact with LQG is the fact that operators associated with triangle areas have a continuous spectrum in the Lorentzian models which is in conflict with the canonical result (7). From this perspective one is compelled to study the model we have described in section 9.3.3 for which part of the area spectrum coincides with (7).

In the canonical framework the compact $SU(2)$ formulation is achieved by means of the introduction of the Immirzi parameter ι which defines a one-parameter family of unitarily inequivalent representations of quantum gravity. In particular, geometric operators are modulated by ι (for example, see (7)). Can the spin foam approach say something about ι? There is no conclusive answer to this question so far. The spin foam quantization of generalizations of Plebanski's action including the ι ambiguity [127] are analysed in [128, 129].

So far we have only compared spin foams with loop quantum gravity at the kinematical level. However, one of the main motivations of the approach was the expected simplification in addressing dynamical questions. As we mentioned before, the dynamics is not well understood in the canonical formulation. This is mostly due to the difficulties associated with regularization ambiguities in the construction of the scalar constraint, the understanding of its space of solutions and the definition of the physical scalar product. As we have seen all these issues can be naturally addressed in the spin foam formulation. In turn, it would be nice to understand whether the approach can help set a guiding principle that would allow for the selection of a preferred regularization of the scalar constraint. Understanding of this question would strengthen the relationship between the two formulations. An explicit analysis of the relation between the covariant (spin foam) and canonical formulation of BF theory is presented in [130]. The reconstruction of the corresponding scalar constraint operator is studied and carried out explicitly in simple models.

4. *Alternatives to the Barrett–Crane model*. The fact that degenerate configurations dominate the asymptotics of the Barrett–Crane vertex amplitude [107] can be interpreted as a serious problem. A possible solution to the apparent problem is proposed in [131], where the evaluation of the vertex amplitude is modified in a simple way to avoid degenerate contributions. The potential dominance of degenerate sectors of Plebanski's real formulation over the gravity sector in a formal path integral has been emphasized on general grounds by Reisenberger in [132]. There is so far no rigorous result linking the previous two. The spin foam quantization of degenerate sectors ($e = 0$) of section 7.10 is an attempt to investigate this problem. A puzzling feature of the degenerate model is that, even when the vertex and face amplitudes are different, it fully contains the Barrett–Crane 4-simplex configurations. Is this an indication of the apparent problem raised by Reisenberger? This motivates the question of whether e in (72) is well defined in the Barrett–Crane model, and if so, whether it differs from zero. The vanishing of e would appear as a serious obstacle to reproducing the gravity sector of Plebanski's theory in the low-energy limit.
5. *Coupling with matter*. The inclusion of matter into the spin foam formalism is of clear importance. Here we review recent results obtained in this direction. A natural generalization would be gravity with cosmological constant. The definition of the analogue of the Barrett–Crane model with cosmological constant has been explored by Roche and Noui in [133]. The presence of the cosmological 'matter' is modelled using the quantum deformation $U_q(SL(2,\mathbb{C}))$ whose representation theory is well understood [134] and the deformation parameter is determined by the value of the cosmological constant. In this modification the homogeneous space H^+ is replaced by the quantum deformation of the hyperboloid defined in terms of a pile of *fuzzy* spheres. In a recent

paper Oriti and Pfeiffer [135] proposed a model which couples the Riemannian Barrett–Crane model with Yang–Mills theory. Their construction is likely to be generalizable to other models. In the context of the GFT formulation of spin foam models Mikovic proposes a way to include matter by adding certain *matter* fields into the GFT-action. In [136, 137] he generalizes the GFT construction described in section 8 by allowing for the inclusion of spinor fields in finite-dimensional representations of $SO(2)$ and $SO(3)$ representing matter corresponding to fermions and gauge fields. States in the theory are given by spin networks with open links; this is consistent with the results obtained in the canonical approach [7, 138–140].

More radical and very appealing possibilities are suggested by Crane. In [141] he proposes a topological QFT of the type of BF theory as fundamental theory. The gravitational degrees of freedom are represented by the subset of (constrained) representations while those of matter are encoded in the remaining ones. The full spin foam model is topological (no local degrees of freedom). In order to recover the low-energy world (with local excitations) the author appeals to certain 'symmetry breaking' of topological invariance. Crane's other proposal consists of interpreting topological (conical) singularities naturally arising in the structure of the Feynman diagrams of the GFT theory as representing matter degrees of freedom [142]. These possibilities are very attractive for their purely geometric character.

6. *Discretization dependence*. In section 4 we have seen that the discretization dependence is trivial in three dimensions. In particular, this has been nicely formalized in the definition of continuum spin foams by Zapata [70]. Some definition of the refinement limit, discussed in section 6.2, should be investigated for the models which are not trivially discretization independent. The model defined in section 7.10 (corresponding to the degenerate sectors of Plebanski's theory) might be an interesting candidate. The model is not topological and has a clear connection to a continuum action. It is somehow between the theory we want to define and the simpler theories we understand well but that do not have local excitations (such as BF theory and three-dimensional gravity). From this viewpoint we believe that it might be useful to explore its behaviour under refinement as a 'toy model'. The simplicity of the model might even allow for analytic computations. Incidentally other open issues in the spin foam approach, such as the problem of the continuum limit, gauge and anomaly questions and the construction of the generalized projector into the physical Hilbert space, should be investigated in this model.

Another proposal for a discretization-independent formulation is given by the GFT formulation. The GFT formulation is very attractive since it provides a discretization-independent formulation of spin foam models from the outset. Also it has been very useful for the definition of the Lorentzian models of sections 9.2 and 9.3 as a device for formal manipulations. However, the mathematical consistency of this definition depends on whether one can make sense of the expansion in λ of equation (105). Suggestions on how the series could be summable by complexification of the coupling λ can be found in [16]. A beautiful example showing that the λ-series can be summed in certain cases has been constructed by Freidel and Louapre [89] for certain three-dimensional GFTs.

In the standard background-dependent realm, the problem of removing the regulator introduced by the discretization is referred to as the *continuum limit*. Note that the nature of the regulator in spin foams is totally different. The discretization introduces a cut-off in the allowed configurations, but the discrete and combinatorial nature of the latter is expected to be preserved when the regulator is removed. When the regulator is removed we expect to obtain a definition of the generalized projection operator which allows for the computation of physical transition amplitudes between spin network states. On the other hand, physical states are expected to lie outside the kinematical Hilbert space and so they could be given by elements in the dual that resemble (in some sense) continuous configurations. In order to avoid confusion about the matter we have consistently referred to the notion of *removal of the discretization dependence* instead of that of the *continuum limit*.

The approach of dynamical triangulations is a background-independent formulation of quantum gravity in which the removal of the regulator is similar in spirit to both the problem in the spin foam context and that of (the background-dependent) lattice gauge theory. Here we present an account of the recent results and point out the conceptual differences with the spin foam approach. We hope that, despite the differences, these results might be helpful in developing useful techniques in the spin foam approach. In dynamical triangulations the nature of the regulator is different from that of spin foam models. In this approach, (diff-equivalent classes of) smooth metric configurations are approximated by spacetime triangulations where 1-simplexes have the same fixed proper length ℓ. The smaller the length scale the better the approximation; therefore, the proper length represents the regulator in the theory to be removed in a certain $\ell \to 0$ limit. The phrase *continuum limit* certainly has a clear-cut meaning here. The continuum limit in dynamical triangulations has been extensively studied. As we previously mentioned, dynamical triangulation is a definition of quantum gravity based on the Euclidean path integral, i.e., configurations are weighted with real amplitudes $\exp(-S)$. In order to recover Lorentzian quantum gravity one has to define what is meant by a Wick rotation in the background-independent framework. Once this is done, the continuum limit can be studied using standard techniques of statistical mechanics. Results indicate that in Euclidean dynamical triangulations there is no continuum limit [143]. The path integral is dominated by singular Euclidean geometries. The situation improves in the Lorentzian models. This is because of the restrictions imposed by the notion of Wick rotation defined. These restrictions are such that in 1+1 gravity the model is exactly solvable and possesses a continuum limit. There is some numerical evidence that could be the case in higher dimensions [144, 145].

7. *Contact with the low-energy world*. In this topical review, we have concentrated mostly on mathematical and conceptual issues which are important in the construction of a consistent spin foam model for gravity. However, one of the most pressing questions is whether the spin foam models can reproduce the low-energy world of general relativity. Although this is certainly one of the most important questions, there is unfortunately no conclusive evidence of this at present and most of the work lies ahead of us.

   Some of the models presented in section 5 satisfy some notion of 'naive' classical limit in the sense that they are derived from a continuous action of general relativity. Other models are defined by postulating the fundamental dynamics and using the kinematical structure discovered by the canonical formulation. If any of these models are to be taken as strong candidates for theories of quantum gravity they must be able to reproduce the physics of gravity at low energy and predict the corresponding semiclassical corrections. In this respect, it is still an unsettled issue whether the differences of the various models at the fundamental level lead to different theories or rather are to be expected to disappear in the continuum limit. After all, the latter happens in lattice gauge theory in the renormalization process. Even when progress has been made in this direction, there is, however, no clear-cut formulation of the analogue of renormalization theory for spin foams. The reason for this is the difficulty of applying standard techniques in the background-independent context. In section 6.4 we mentioned some new ideas and attempts to tackle this problem. This is an extremely important problem where new ideas will have to play an important role.

8. *The generalized projection versus the Feynman proper time propagator*. Throughout this work we have advocated the viewpoint of the path integral as a device for the computation of the generalized projection operator into the physical Hilbert space of quantum gravity. We claimed that this picture is forced upon us by the general covariance of general relativity. Although we have kept the term *transition amplitude* for notational convenience there is no notion of *time* involved in our definitions.

   In other path integral approaches one aims at the construction of the notion of transition amplitudes between 3-geometries at a definite *time*. Such transition amplitudes are governed by the so-called proper time evolution operator or proper time propagator.

   This possibility has also been explored in the context of LQG [53] and is the perspective in which the causal models of section 5.5 are defined. There are also modifications of the Barrett–Crane model which attempt to define such a proper time evolution operator [146, 147]. The main difficulty in such an approach is how to give meaning to the notion of 'time' in the general covariant context.

The suggestion that quantum gravity should be described in terms of discrete combinatorial structures can be traced all the way back to Einstein [148]. In this topical review, we have described an approach inspired by the non-perturbative canonical quantization of general relativity in which these suggestions are concretely realized. We have shown that a great deal of progress has been achieved in understanding the conceptual issues involved. The fundamental motivation was the construction of a theory that would provide a device to analyse dynamics in quantum gravity. The conceptual setting is clear: spin foams should provide the definition of the physical scalar product and hence the physical Hilbert space of quantum gravity which encodes all the information about quantum dynamics. We have seen that these models can be obtained as lattice quantizations of general relativity in appropriate variables, the 'lattice' action being naively related to the continuum general relativity action (or some equivalent classical formulation) in the sense of Wilson's action for lattice gauge theory. Some of these models have remarkable finiteness properties both in the Riemannian and Lorentzian sectors. We have also shown how the basic structure of spin foams arises from the canonical formulation of loop quantum gravity. In addition, we described models which are not related to a classical action and are constructed from the basic properties of spin networks plus simple causality requirements. For all these models the common structure arising is given by spin foams: coloured 2-complexes where the geometric degrees of freedom are encoded in a fully combinatorial manner. Spin foam models appear as a beautiful realization of Einstein's idea. There are certainly many difficult open questions and we have tried to point out those which we judge the most important. We hope that new ideas and hard work will continue to contribute to their resolution in the near future.

Acknowledgments

I thank Abhay Ashtekar, John Baez, Martin Bojowald, Dan Christensen, Rodolfo Gambini, Amit Ghosh, Jorge Pullin, Michael Reisenberger, Carlo Rovelli and Lee Smolin for discussions. I thank Martin and Carlo for suggestions that helped improve the presentation of this work. I am grateful to Florian Girelli, Josh Willis and Jacek Wisniewski for their careful reading of the manuscript. I specially thank Carlo Rovelli for his strong support. This work was supported in part by NSF Grant PHY-0090091, and Eberly Research Funds of Penn State.

References

[1] Rovelli C 2000 Notes for a brief history of quantum gravity *Preprint* gr-qc/0006061

[2] Weinberg S 1995 *The Quantum Theory of Fields* (Cambridge: Cambridge University Press)

[3] Ashtekar A and Lewandowski J 2003 Background independent quantum gravity: a status report *Preprint* to appear

[4] Thieman T 2001 Introduction to modern canonical quantum general relativity *Preprint* gr-qc/0110034

[5] Gaul M and Rovelli C 2000 Loop quantum gravity and the meaning of diffeomorphism invariance *Lect. Notes Phys.* **541** 277–324

[6] Rovelli C 1998 Loop quantum gravity *Living Rev. Gen. Rel.* **1**

[7] Thieman T 1998 Quantum gravity as the natural regulator of matter quantum field theories *Class. Quantum Grav.* **15** 1281–314

[8] Thiemann T 2001 Gauge field theory coherent states (GCS): I. General properties *Class. Quantum Grav.* **18** 2025

[9] Barrett J W and Crane L 1998 Relativistic spin networks and quantum gravity *J. Math. Phys.* **39** 3296–302

[10] Barrett J W and Crane L 2000 A Lorentzian signature model for quantum general relativity *Class. Quantum Grav.* **17** 3101–18

[11] De Pietri R, Freidel L, Krasnov K and Rovelli C 2000 Barrett–Crane model from a Boulatov–Ooguri field theory over a homogeneous space *Nucl. Phys.* B **574** 785–806

[12] Perez A and Rovelli C 2001 A spin foam model without bubble divergences *Nucl. Phys.* B **599** 255–82

[13] Perez A and Rovelli C 2001 Spin foam model for Lorentzian general relativity *Phys. Rev.* D **63** 041501

[14] Baez J C 1998 Spin foam models *Class. Quantum Grav.* **15** 1827–58

[15] Baez J C and Barrett J W 1999 The quantum tetrahedron in 3 and 4 dimensions *Adv. Theor. Math. Phys.* **3** 815–50

[16] Reisenberger M P and Rovelli C 2001 Spacetime as a Feynman diagram: the connection formulation *Class. Quantum Grav.* **18** 121–40

[17] Reisenberger M P and Rovelli C 2000 Spin foams as Feynman diagrams *Preprint* gr-qc/0002083

[18] Perez A 2001 Finiteness of a spinfoam model for Euclidean quantum general relativity *Nucl. Phys.* B **599** 427–34

[19] Crane L, Perez A and Rovelli C 2001 A finiteness proof for the Lorentzian state sum spin foam model for quantum general relativity *Preprint* gr-qc/0104057

[20] Crane L, Perez A and Rovelli C 2001 Perturbative finiteness in spin-foam quantum gravity *Phys. Rev. Lett.* **87** 181301

[21] Ashtekar A 1991 *Lectures on Non Perturbative Canonical Gravity* (Singapore: Word Scientific)

[22] Ashtekar A 1986 New variables for classical and quantum gravity *Phys. Rev. Lett.* **57** 2244

[23] Barbero J F 1995 Real Ashtekar variables for Lorentzian signature space times *Phys. Rev.* D **51** 5507–10

[24] Ashtekar A and Lewandowski J 1995 Projective techniques and functional integration *J. Math. Phys.* **36** 2170

[25] Reisenberger M 1994 World sheet formulations of gauge theories and gravity *Preprint* gr-qc/941235

[26] Rovelli C and Smolin L 1995 Spin networks and quantum gravity *Phys. Rev.* D **53** 5743

[27] Baez J 1996 Spin network states in gauge theory *Adv. Math.* **117** 253–72

[28] Penrose R 1971 *Quantum Theory and Beyond* ed T Bastin (Cambridge: Cambridge University Press)

[29] Kogut J B and Susskind L 1975 Hamiltonian formulation of Wilson's and lattice gauge theories *Phys. Rev.* D **11** 395

[30] Furmanski W and Kolawa A 1987 Yang–Mills vacuum: an attempt of lattice loop calculus *Nucl. Phys.* B **291** 594

[31] Smolin L 1997 The future of spin networks *Preprint* gr-qc/9702030

[32] Zapata J A 1999 A combinatorial approach to quantum gauge theories and quantum gravity *Preprint* UMI-99-01167

[33] Smolin L 1992 Recent developments in non perturbative quantum gravity *Quantum Gravity and Cosmology, Proc. 1991 GIFT Int. Seminar on Theoretical Physics* (Singapore: World Scientific) webpage hep-th/9202022

[34] Rovelli C and Smolin L 1995 Discreteness of the area and volume in quantum gravity *Nucl. Phys.* B **442** 593
Rovelli C and Smolin L 1995 *Nucl. Phys.* B **456** 734 (erratum)

[35] Ashtekar A and Lewandowski J 1997 Quantum theory of gravity: I. Area operators *Class. Quantum Grav.* **14** A55–A81

[36] Immirzi G 1997 Real and complex connections for canonical gravity *Class. Quantum Grav.* **14** L177–81

[37] Ashtekar A, Baez J C and Krasnov K 2000 Quantum geometry of isolated horizons and black hole entropy *Adv. Theor. Math. Phys.* **4** 1–94

[38] Loll R 1997 Simplifying the spectral analysis of the volume operator *Nucl. Phys.* B **500** 405–20
[39] Ashtekar A and Lewandowski J 1997 Quantum theory of gravity: II. Volume operators *Preprint* gr-qc/9711031
[40] Thieman T 1996 Anomaly-free formulation of non perturbative, four dimensional Lorentzian quantum gravity *Phys. Lett.* B **380** 257
[41] Gaul M and Rovelli C 2001 A generalized Hamiltonian constraint operator in loop quantum gravity and its simplest Euclidean matrix elements *Class. Quantum Grav.* **18** 1593–624
[42] Di Bartolo C, Gambini R, Griego J and Pullin J 2000 Canonical quantum gravity in the Vassiliev invariants arena: II. Constraints, habitats and consistency of the constraint algebra *Class. Quantum Grav.* **17** 3239–64
[43] Di Bartolo C, Gambini R, Griego J and Pullin J 2000 Consistent canonical quantization of general relativity in the space of Vassiliev knot invariants *Phys. Rev. Lett.* **84** 2314–7
[44] Halliwell J and Hartle J B 1991 Wave functions constructed from an invariant sum over histories satisfy constraints *Phys. Rev.* D **43** 1170–94
[45] Reisenberger M and Rovelli C 2002 Spacetime states and covariant quantum theory *Phys. Rev.* D **65** 125016
[46] Jacobson T and Smolin L 1988 Non perturbative quantum geometries *Nucl. Phys.* B **299** 295
[47] Smolin L and Rovelli C 1990 Loop space representation of quantum general relativity *Nucl. Phys.* B **331** 80
[48] Hartle J B and Hawking S W 1983 Wave function of the universe *Phys. Rev.* D **28** 2960–75
[49] Gibbons G W and Hawking S W 1993 *Euclidean Quantum Gravity* (Singapore: World Scientific) p 586
[50] Crane L 1993 Categorical physics *Preprint* hep-th/9301061
[51] Crane L 1993 Topological field theory as the key to quantum gravity *Preprint* hep-th/9308126
[52] Crane L and Yetter D 1993 A categorical construction of 4-D topological quantum field theories *Quantum Topology* ed L Kaufmann and R Baadhio (Singapore: World Scientific)
[53] Reisenberger M P and Rovelli C 1997 'Sum over surfaces' form of loop quantum gravity *Phys. Rev.* D **56** 3490–508
[54] Rovelli C 1999 The projector on physical states in loop quantum gravity *Phys. Rev.* D **59** 104015
[55] Thieman T 1998 Kinematical Hilbert spaces for fermionic and Higgs quantum field theories *Class. Quantum Grav.* **15** 1487–512
[56] Reisenberger M P 1997 A left-handed simplicial action for Euclidean general relativity *Class. Quantum Grav.* **14** 1753–70
[57] Reisenberger M P 1997 A lattice worldsheet sum for 4-d Euclidean general relativity *Preprint* gr-qc/9711052
[58] Plebanski J F 1977 On the separation of Einsteinian substructures *J. Math. Phys.* **18** 2511
[59] Baez J C 2000 An introduction to spin foam models of quantum gravity and BF theory *Lect. Notes Phys.* **543** 25–94
[60] Freidel L 2000 A Ponzano–Regge model of Lorentzian 3-dimensional gravity *Nucl. Phys. Proc. Suppl.* **88** 237–40
[61] Iwasaki J 1995 A definition of the Ponzano–Regge quantum gravity model in terms of surfaces *J. Math. Phys.* **36** 6288–98
[62] Iwasaki J 1994 A reformulation of the Ponzano–Regge quantum gravity model in terms of surfaces *Preprint* gr-qc/9410010
[63] Vilenkin N and Klimyk A U 1993 *Representation of Lie Groups and Special Functions* (Dordrecht: Kluwer)
[64] Ponzano G and Regge T 1968 Semiclassical limit of Racah coefficients *Spectroscopy and Group Theoretical Methods in Physics* ed F Block *et al* (Amsterdam: North-Holland)
[65] Freidel L and Louapre D 2002 Diffeomorphisms and spin foam models *Preprint* gr-qc/0212001
[66] Turaev V G and Viro O Y 1992 Statesum invariants of 3-manifolds and quantum 6*j*-symbols *Topology* **31** 865–902
[67] Turaev V G 1994 *Quantum Invariants of Knots and 3-Manifolds* (Berlin: de Gruyter)
[68] Kauffman L H and Lins S L 1994 *Temperley–Lieb Recoupling Theory and Invariants of 3-Manifolds* (Princeton, NJ: Princeton University Press)
[69] Girelli F, Oeckl R and Perez A 2002 Spin foam diagrammatics and topological invariance *Class. Quantum Grav.* **19** 1093–108
[70] Zapata J A 2002 Continuum spin foam model for 3d gravity *Preprint* gr-qc/0205037
[71] Freidel L and Krasnov K 1999 Spin foam models and the classical action principle *Adv. Theor. Math. Phys.* **2** 1183–247
[72] Iwasaki J 1999 A surface theoretic model of quantum gravity *Preprint* gr-qc/9903112
[73] Iwasaki J 2000 A lattice quantum gravity model with surface-like excitations in 4-dimensional spacetime *Preprint* gr-qc/0006088
[74] Barbieri A 1997 Quantum tetrahedra and simplicial spin networks *Preprint* gr-qc/9707010
[75] Markopoulou F and Smolin L 1997 Causal evolution of spin networks *Nucl. Phys.* B **508** 409–30
[76] Markopoulou F 1997 Dual formulation of spin network evolution *Preprint* gr-qc/9704013
[77] Markopoulou F and Smolin L 1998 Quantum geometry with intrinsic local causality *Phys. Rev.* D **58** 084032

[78] Gambini R and Pullin J 2001 A finite spin-foam-based theory of three and four dimensional quantum gravity *Preprint* gr-qc/0111089

[79] Thiemann T 1998 QSD IV: 2 + 1 Euclidean quantum gravity as a model to test 3 + 1 Lorentzian quantum gravity *Class. Quantum Grav.* **15** 1249–80

[80] Capovilla R, Jacobson T and Dell J 1991 A pure spin connection formulation of gravity *Class. Quantum Grav.* **8** 59–73

[81] Capovilla R, Jacobson T and Dell J 1989 General relativity without the metric *Phys. Rev. Lett.* **63** 2325

[82] Freidel L and Livine E R 2002 Spin networks for non compact groups *Preprint* hep-th/0205268

[83] Gambini R and Pullin J 2002 Canonical quantization of general relativity in discrete spacetimes *Preprint* gr-qc/0206055

[84] Di Bartolo C, Gambini R and Pullin J 2002 Canonical quantization of constrained theories on discrete spacetime lattices *Preprint* gr-qc/0206055

[85] Bojowald M and Perez A spin foam quantization and anomalies (in preparation)

[86] Boulatov D 1992 A model of three-dimensional lattice gravity *Mod. Phys. Lett.* A **7** 1629–46

[87] Ooguri H 1992 Topological lattice models in four dimensions *Mod. Phys. Lett.* A **7** 2799–810

[88] Petronio C and De Pietri R 2000 Feynman diagrams of generalized matrix models and the associated manifolds in dimension 4 *J. Math. Phys.* **41** 6671–88

[89] Freidel L and Louapre D 2002 Non-perturbative summation over 3D discrete topologies *Preprint* hep-th/0211026

[90] Livine R, Perez A and Rovelli C 2001 2d manifold-independent spinfoam theory *Preprint* gr-qc/0102051

[91] Haag R 1996 *Local Quantum Physics: Fields, Particles, Algebras* (New York: Springer)

[92] Perez A and Rovelli C 2001 Observables in quantum gravity *Preprint* gr-qc/0104034

[93] Baez J C, Christensen J D, Halford T R and Tsang D C 2002 Spin foam models of Riemannian quantum gravity *Preprint* gr-qc/0202017

[94] Markopoulou F 2002 Coarse graining in spin foam models *Preprint* gr-qc/0203036

[95] Markopoulou F 2000 An algebraic approach to coarse graining *Preprint* hep-th/0006199

[96] Kreimer D 1998 On the Hopf algebra structure of perturbative quantum field theories *Adv. Theor. Math. Phys.* **2** 303–34

[97] Kreimer D 1999 On overlapping divergences *Commun. Math. Phys.* **204** 669

[98] Oeckl R 2002 Renormalization of discrete models without background *Preprint* gr-qc/0212047

[99] Perez A 2001 Spinfoam models for quantum gravity *PhD Thesis* University of Pittsburgh, Pittsburgh

[100] Crane L, Kauffman L and Yetter D N 1997 State-sum invariants of 4-manifolds *J. Knot Theor. Ramifications* **6** 177–234

[101] De Pietri R and Freidel L 1999 *so*(4) Plebanski action and relativistic spin foam model *Class. Quantum Grav.* **16** 2187–96

[102] Reisenberger M P 1999 On relativistic spin network vertices *J. Math. Phys.* **40** 2046–54

[103] Barrett J W 1998 The classical evaluation of relativistic spin networks *Adv. Theor. Math. Phys.* **2** 593–600

[104] Major S A and Seifert M D 2002 Modelling space with an atom of quantum geometry *Class. Quantum Grav.* **19** 2211–28

[105] Crane L and Yetter D N 1994 On the classical limit of the balanced state sum *Preprint* gr-qc/9712087

[106] Barrett J W and Williams R M 1999 The asymptotics of an amplitude for the 4-simplex *Adv. Theor. Math. Phys.* **3** 209–15

[107] Baez J C, Christensen J D and Egan G 2002 Asymptotics of 10*j* symbols *Preprint* gr-qc/0208010

[108] Freidel L and Louapre D 2002 Asymptotics of 6*j* and 1D*j* symbols *Preprint* hep-th/0209134

[109] Barrett J W and Steele C M 2002 Asymptotics of relativistic spin networks *Preprint* gr-qc/0209023

[110] Perez A and Rovelli C 2001 3 + 1 spinfoam model of quantum gravity with spacelike and timelike components *Phys. Rev.* D **64** 064002

[111] Baez J C and Christensen J D 2002 Positivity of spin foam amplitudes *Class. Quantum Grav.* **19** 2291–306

[112] Pfeiffer H 2002 Positivity of relativistic spin network evaluations *Preprint* gr-qc/0211106

[113] Mikovic A 2001 Quantum field theory of spin networks *Class. Quantum Grav.* **18** 2827–50

[114] Wightman A S 1956 Quantum field theory in terms of vacuum expectation values *Phys. Rev.* **101** 860

[115] Wightman A S 1964 *PCT, Spin Statistics and All That* (New York: Benjamin–Cummings)

[116] Oriti D and Williams R M 2001 Gluing 4-simplices: a derivation of the Barrett–Crane spin foam model for Euclidean quantum gravity *Phys. Rev.* D **63** 024022

[117] Gel'fand I M 1966 *Generalized Functions* vol 5 (New York: Academic)

[118] Baez J C and Barrett J W 2001 Integrability for relativistic spin networks *Class. Quantum Grav.* **18** 4683–700

[119] Oriti D 2002 Boundary terms in the Barrett–Crane spin foam model and consistent gluing *Phys. Lett.* B **532** 363–72

[120] Pfeiffer H 2002 Dual variables and a connection picture for the Euclidean Barrett–Crane model *Class. Quantum Grav.* **19** 1109–38 (*Preprint* gr-qc/0112002)
[121] Christensen D, Daniel J and Egan G 2002 An efficient algorithm for the Riemannian 10*j* symbols *Class. Quantum Grav.* **19** 1185–94
[122] Smolin L 2002 Private communication
[123] Alexandrov S 2002 Hilbert space structure of covariant loop quantum gravity *Phys. Rev.* D **66** 024028
[124] Alexandrov S 2002 Choice of connection in loop quantum gravity *Phys. Rev.* D **65** 024011
[125] Alexandrov S and Livine E R 2002 *SU*(2) loop quantum gravity seen from covariant theory *Preprint* gr-qc/0209105
[126] Livine E R 2002 Projected spin networks for Lorentz connection: linking spin foams and loop gravity *Class. Quantum Grav.* **19** 5525–42 (*Preprint* gr-qc/0207084)
[127] Capovilla R, Montesinos M, Prieto V A and Rojas E 2001 BF gravity and the Immirzi parameter *Class. Quantum Grav.* **18** L49
[128] Livine R E and Oriti D 2002 Barrett–Crane spin foam model from generalized Bf-type action for gravity *Phys. Rev.* D **65** 044025
[129] Livine R E 2001 Immirzi parameter in the Barrett–Crane Model *Preprint* gr-qc/0103081
[130] Arnsdorf M 2002 Relating covariant and canonical approaches to triangulated models of quantum gravity *Class. Quantum Grav.* **19** 1065–92
[131] Freidel L and Louapre D 2002 Asymptotics of 6*j* and 10*j* symbols *Preprint* hep-th/0209134
[132] Reisenberger M P 1998 Classical Euclidean general relativity from 'left-handed area = right-handed area' *Preprint* gr-qc/9804061
[133] Noui K and Roche Ph 2002 Cosmological deformation of Lorentzian spin foam models *Preprint* gr-qc/0211109
[134] Buffenoir E and Roche P 1999 Harmonic analysis on the quantum Lorentz group *Commun. Math. Phys.* **207** 499–555
[135] Oriti D and Pfeiffer H 2002 A spin foam model for pure gauge theory coupled to quantum gravity *Preprint* gr-qc/0207041
[136] Mikovic A 2002 Spin foam models of matter coupled to gravity *Class. Quantum Grav.* **19** 2335–54
[137] Mikovic A 2002 Quantum field theory of open spin networks and new spin foam models *Preprint* gr-qc/0202026
[138] Morales-Tecotl H A and Rovelli C 1995 Loop space representation of quantum fermions and gravity *Nucl. Phys.* B **451** 325–61
[139] Morales-Tecotl H A and Rovelli C 1994 Fermions in quantum gravity *Phys. Rev. Lett.* **72** 3642–5
[140] Baez J and Krasnov K 1998 Quantization of diffeomorphism-invariant theories with fermions *J. Math. Phys.* **39** 1251–71
[141] Crane L 2002 Hypergravity and categorical Feynmanology *Preprint* gr-qc/0004043
[142] Crane L 2001 A new approach to the geometrization of matter *Preprint* gr-qc/0110060
[143] Ambjorn J 2002 Simplicial Euclidean and Lorentzian quantum gravity *Preprint* gr-qc/0201028
[144] Ambjorn J, Jurkiewicz J and Loll R 2000 A non perturbative Lorentzian path integral for gravity *Phys. Rev. Lett.* **85** 924–7
[145] Ambjorn J, Jurkiewicz J and Loll R 2002 3d Lorentzian, dynamically triangulated quantum gravity *Nucl. Phys. Proc. Suppl.* **106** 980–2
[146] Livine E R and Oriti D 2002 Implementing causality in the spin foam quantum geometry *Preprint* gr-qc/0210064
[147] Pfeiffer H 2002 On the causal Barrett–Crane model: measure, coupling constant, Wick rotation, symmetries and observables *Preprint* gr-qc/0212049
[148] Stachel J 1991 Einstein and quantum mechanics *Conceptual Problems of Quantum Gravity* ed A Ashtekar and J Stachel (Boston, MA: Birkhauser)