In quantum gravity, summing is refining

Carlo Rovelli and Matteo Smerlak

Centre de Physique Théorique, Campus de Luminy, Case 907, 13288 Marseille Cedex 09, France
E-mail: rovelli@cpt.univ-mrs.fr and smerlak@cpt.univ-mrs.fr

Received 23 November 2011, in final form 23 November 2011
Published 9 February 2012
Online at stacks.iop.org/CQG/29/055004

Abstract

In perturbative QED, the approximation is improved by summing more Feynman graphs, while in non-perturbative QCD, by refining the lattice. Here we observe that in quantum gravity, the two procedures may well be the same. We outline the combinatorial structure of spinfoam quantum gravity, define the continuum limit and show that under general conditions, refining foams is the same as summing over them. The conditions bear on the cylindrical consistency of the spinfoam amplitudes and on the presence of appropriate combinatorial factors, related to the implementation of diffeomorphism invariance. Intuitively, the sites of the lattice are points of space: these are themselves quanta of the gravitational field, and thus a lattice discretization is also a Feynman history of quanta.

PACS numbers: 04.60.Pp, 04.60.-m

A key step in constructing an interactive quantum field theory (QFT) is always a finite truncation of the dynamical degrees of freedom. In weakly coupled theories, such as low-energy QED or high-energy QCD, one can rely on the particle structure of the free field and consider virtual processes involving finitely many particles, described by Feynman diagrams. In strongly coupled theories, such as confining QCD, we must resort to a non-perturbative truncation, such as a finite lattice approximation. In either case, the relevant effect of the remaining degrees of freedom can be subsumed under a dressing of the coupling constants, if a criterion of renormalizability or criticality is met. The full theory is then formally defined by a limit where all the degrees of freedom are recovered. The ways in which this limit is implemented in perturbative QED and in lattice QCD, however, are quite different: in the first case, by summing over an increasing number of Feynman graphs; in the second case, by refining the lattice. Here we show that in quantum gravity, the two implementations of the limits may in fact turn out to be aspects of the same structure. In conclusion, we present an intuitive physical interpretation of this result.

It is well known that neither the QED nor the lattice QCD scheme is available as such in quantum gravity: the perturbative expansion is non-renormalizable, and there is no background metric space to discretize. But we can still introduce a truncation of the dynamical degrees of freedom. For instance, in dynamical-triangulation approaches [1], discretized metric degrees
of freedom are captured by fixing edge lengths of the simplices and varying their number and adjacency relations, while in Regge calculus [2], we fix the triangulation and vary the edge lengths. In the first case, the full dynamics is recovered by summing over all triangulations, like one sums over Feynman graphs in perturbative QFT; in the second case, by the limit of infinite refinement of the triangulation, like one refines the lattice in defining the continuum limit of lattice QFT. In loop quantum gravity (LQG), on the other hand, the discretization is induced by the diffeomorphism-invariant kinematics of gravity itself [3]: the quantum states of space are labeled by 3-diffeomorphism classes of spin networks, which are discrete structures in themselves [4]. The spinfoam formalism defines truncated transition amplitudes between two such states [5]. The amplitude is associated with a finite 2-complex, or ‘foam’, cobording the corresponding graphs and colored with irreducible representations of SU(2) (for a recent introduction to the covariant theory, see [6]). The question of how to recover the full set of degrees of freedom in this context has been much debated. Should we sum over all foams, or refine them and take a continuum limit? Are spinfoams like Feynman diagrams and dynamical triangulations, or like metric lattices and Regge manifolds? We show below that, under suitable conditions, the two alternatives are in fact the same: the amplitude of a foam is equal to the sum of the amplitudes of all its subfoams, with the trivial representation of SU(2) excluded from the admissible colorings.

The conditions for this to happen are of interest. First, the vertex amplitude must obey a ‘cylindrical consistency’ requirement. The Euclidean vertex amplitude introduced in [7] and [8] satisfies this condition, and so does the Lorentzian amplitude (but not the variant proposed in [9]). Related conditions are studied by Magliaro and Perini [10] and by Bahr, Hellmann, Kaminski, Kisiełowski and Lewandowski [11]. Second, suitable combinatorial factors must be included in the definition of the amplitude. These have a compelling interpretation: they correspond to quotienting by the volume of the residual discrete action of the diffeomorphism group on foams.

Spinfoam models for quantum gravity can be seen as combinatorial relatives of topological quantum field theories (TQFT) [5]. In Atiyah’s scheme, an \((n + 1)\)-dimensional TQFT is as a functorial association of a finite-dimensional Hilbert space \(\mathcal{H}_\Sigma\) with each closed oriented \(n\)-manifold \(\Sigma\), and a vector \(Z_M \in \mathcal{H}_\Sigma\) with each oriented \((n + 1)\)-manifold \(M\) having \(\Sigma\) as its boundary [12, 13]. Similarly, a spinfoam model associates with each oriented graph \(\Gamma\) a Hilbert space \(\mathcal{H}_\Gamma\), and with each foam \(C\) having \(\Gamma\) as its boundary a vector \(Z_C \in \mathcal{H}_\Gamma\). In order to make this analogy precise, let us start with some relevant definitions. (Graphs are topological spaces as well as combinatorial objects. Likewise, foams can be defined in two ways, emphasizing their topological or their combinatorial structure. Although the former approach is often taken in the literature, after Baez [14, 15], we find it more convenient to focus on the relevant combinatorics from scratch.)

We call graph a couple \(\Gamma = (N_\Gamma, L_\Gamma)\) with \(N_\Gamma\) a finite set of nodes and \(L_\Gamma\) a set of ordered pairs of nodes, the links of \(\Gamma\). The nodes \(n\) and \(n'\) of a link \(l = (n, n')\) are called the source and target of \(l\), and denoted by \(s(l)\) and \(t(l)\), respectively. We denote by \(t^{-1} = (n', n)\) the reversed link and by \(\Gamma\) the graph obtained from \(\Gamma\) reversing all links.

A combinatorial 2-complex, or foam, is a triple \(C = (V_C, E_C, F_C)\) with \(V_C\) a finite set of vertices, \(E_C\) a set of ordered pairs of vertices \(e = (v, v')\), now called edges, and \(F_C\) a finite set of faces. Here, we call face a finite sequence of edges \(f = (e_1, \ldots, e_t)\), where \(t(e_{n+1}) = s(e_n)\), \(\epsilon_n = \pm 1\) and \(n_{\text{e}} + 1 := 1\). Note that any subset \(F\) of \(F_C\) naturally defines a subcomplex of \(C\), made of the vertices, edges and faces appearing in \(F\). Finally, we denote by \(\tilde{C}\) the foam obtained from \(C\) by reversing all its edges and faces.

There is a notion of boundary for foams. Indeed, let us call the edges of \(E_C\) appearing exactly once in only one face its links, and the other ones its interior edges. Similarly, let us
call the vertices appearing exactly once in an interior edge its **nodes**, and the other ones its **interior vertices**. The sets of nodes and links of a foam \( C \) generally do not form a graph, but when they do, and moreover when the orientation of each link matches the one induced by the unique face passing through it, we say that \( C \) is a **proper foam**. We then define the boundary \( \partial C \) as the subcomplex of \( C \) defined by those faces of \( C \) which contain at least one link. The underlying graph of \( \partial C \) is the **boundary graph** of \( C \). In this language, foams with the same boundary graphs can still have different boundaries, because the boundary faces can be different.

We say that two proper foams \( C_1 \) and \( C_2 \) are **composable** along an oriented graph \( \Gamma \) if \( \Gamma \) is a connected component of the boundary graphs of both \( C_1 \) and \( C_2 \). In this case, we can define the **composition** of \( C_1 \) and \( C_2 \) along \( \Gamma \) as the foam \( C_1 \cup \Gamma C_2 \) obtained by removing \( \Gamma \) and merging, for each node \( n \), the unique edges \( e_n^1 \) and \( e_n^2 \) of \( C_1 \) and \( C_2 \) adjacent on \( n \) into a single edge, and for each link \( l \) the unique faces \( f_l^1 \) and \( f_l^2 \) of \( C_1 \) and \( C_2 \) adjacent on \( l \) into a single face (see figure 1). Note that there is an obvious unit \( \mathbb{1}_\Gamma \) for the composition along \( \Gamma \), the ‘cylinder’ with exactly one face per link of \( \Gamma \).

With these definitions, we can define a general spinfoam model in the manner of Atiyah as the association of a Hilbert space \( \mathcal{H}_\Gamma \) with each oriented graph \( \Gamma \), and of a vector \( Z_C \in \mathcal{H}_\Gamma \) with each proper foam \( C \) with the boundary graph \( \Gamma \). This association should satisfy the following axioms:

- (multiplicativity) \( \mathcal{H}_{\Gamma_1 \cup \Gamma_2} = \mathcal{H}_{\Gamma_1} \otimes \mathcal{H}_{\Gamma_2} \),
- (duality) \( \mathcal{H}_\emptyset = \mathcal{H}_\emptyset^* \) and \( Z_{\emptyset} = Z_{\emptyset}^* \),
- (functoriality) \( Z_{C_1 \cup \Gamma C_2} = (Z_{C_1} | Z_{C_2})_{\mathcal{H}_\Gamma} = (Z_{C_1} | Z_{C_2})_{\mathcal{H}_\Gamma} \), to which we add the obvious axioms \( \mathcal{H}_\emptyset = \mathbb{C} \) and \( Z_{\emptyset} = \text{id}_{\mathcal{H}_\emptyset} \) (using the canonical identification of elements of \( \mathcal{H}_{\Gamma_1} \otimes \mathcal{H}_{\Gamma_2}^* \) with operators on \( \mathcal{H}_\Gamma \)).

This set of axioms is precisely Atiyah’s, except for two points. (See also Oeckl’s **general boundary** formulation of QFT [16].) First, because the boundary graphs are not oriented globally, like boundary manifolds, but only linkwise, the condition of functoriality cannot be framed in terms of morphisms between source and target spaces, but rather in terms of partial inner products as above. Second, we do not require the Hilbert spaces to be finite dimensional.

The dynamics of LQG, as defined in [5, 6], fits this framework. In this case, the Hilbert spaces associated with the graph \( \Gamma \) are the lattice \( SU(2) \) Yang–Mills space \( \mathcal{H}_\Gamma = \mathcal{L}^2(SU(2))^{|\mathcal{L}(\Gamma)|/SU(2)^{|\mathcal{L}(\Gamma)|}} \) (modulo the automorphisms of \( \Gamma \)), and the **foam amplitudes** \( Z_C \) are defined themselves as the sums \( Z_C = \sum_\sigma Z_C(\sigma) \) of the **spinfoam amplitudes** \( Z_C(\sigma) \) over colorings \( \sigma \) of the faces of \( C \) by unitary irreducible representations \( j_f \) of \( SU(2) \) (see below).
From a physical perspective, the vectors $\psi_C \in \mathcal{H}_C$ are quantum states of the spatial geometry, and the inner products $\langle Z_C | \psi_C \rangle$ the quantum amplitude of a spacetime truncated by $C$. To remove the dependence on $C$, two options can be envisaged: infinitely refining $C$ and summing over $C$. Since the set of foams is discrete, the latter is easy to define in principle, at least if one disregards convergence issues. But what about the former?

It can too, thanks to the following observation: the set of proper foams with a given boundary is naturally a directed set, which is a partially ordered set where every pair of elements has an upper bound. Consider indeed the natural order relation defined by $C_1 \leq C_2$ if there is an embedding $i : C_1 \rightarrow C_2$, namely a triple of injective maps $i_v : V_{C_1} \rightarrow V_{C_2}$, $i_e : E_{C_1} \rightarrow E_{C_2}$, $i_f : F_{C_1} \rightarrow F_{C_2}$ preserving the relations between vertices, edges and faces. It is easy to see that any pair of foams with the same boundary has an upper bound $C \geq C_1, C_2$ with respect to this relation, given essentially by attaching the vertices, edges and faces of $C_2$ to $C_1$. Thus, a spinfoam model is a net over the set of proper foams with a given boundary. The continuum limit of $Z_C$, with the fixed boundary $\partial C = c$ and boundary graph $\Gamma$, can then be defined the limit of this net in a suitable topology in $\mathcal{H}_\Gamma$, i.e. the vector $Z_C^\infty \in \mathcal{H}_\Gamma$ such that, for every neighborhood $U$ of $Z_C^\infty$, there is a foam $C_U$ such that $Z_C \in U$ whenever $C \geq C_U$.

What is the relation between the limit $Z_C^\infty$ and the sum $\sum_{C \in \mathbb{C}_c} Z_C$ (assuming that both exist)? First, observe that any choice of a spinfoam amplitude $Z_C(\sigma)$ defines two spinfoam models, depending on whether the trivial representations are included or not:

$$Z_C := \sum_{\sigma \in \text{Col}(C)} Z_C(\sigma), \quad \text{and} \quad Z_C^\infty := \sum_{\sigma \in \text{Col}^\ast(C)} Z_C(\sigma).$$

Here, for a fixed foam $C$, we denoted by Col$(C)$ the set of all colorings $\sigma$ and by Col$^\ast(C)$ the subset consisting of colorings by non-trivial ($j_f \neq 0$) representations.

The relationship between the continuum limit $Z_C^\infty$ and the sum of $Z_C$ over all the foams $C$ with the fixed boundary is as follows: under a general condition on the amplitude $Z_C(\sigma)$, which we detail in the next paragraph, we have

$$Z_C = \sum_{\overset{C \in \mathbb{C}_c}{\delta C = c}} Z_C^\infty,$$  \hspace{1cm} (2)

and hence, passing to the limit (assuming it exists),

$$Z_C^\infty = \sum_{\delta C = c} Z_C^\ast.$$  \hspace{1cm} (3)

This is our central result. It states that refining foams in the model $Z$ is the same as summing over foams in the model $Z^\ast$. Note that, due to the peculiar net structure, all foams appear in (3) in the following sense: every foam $C$ with the boundary $\delta C = c$ appears in one finite sum (2) whose value can be chosen arbitrarily close to $Z_C^\infty$.

The condition on $Z_C(\sigma)$ for relation (3) to hold is one of cylindrical consistencies. First, observe that each coloring $\sigma$ of a foam $C$ comes with a multiplicity, related to the symmetries of $C$. The multiplicity $|\sigma|_C$ of a coloring $\sigma \in \text{Col}(C)$ is the number of colorings $\sigma'$ such that $\sigma = \sigma' \circ \phi$, with $\phi$ an automorphism of $C$, i.e. a bijective embedding of $C$ into itself. Then we say that the amplitude

$$A_C(\sigma) := |\sigma|_C Z_C(\sigma)$$  \hspace{1cm} (4)

is cylindrically consistent if $A_C(\sigma) = A_C(\sigma')$ when $(C', \sigma')$ is a trivial extension of $(C, \sigma)$, that is, when $C$ is a subfoam of $C'$, $\sigma$ and $\sigma'$ coincide on the faces of $C$ and $\sigma'$ is trivial on the other faces of $C'$. 
Under this condition, the proof of (2) is straightforward. First, observe that the subfoams of \( C \) index a partition of \( \text{Col}(C) \), in which each class is made of the trivial extensions of a given subfoam \( C' \subset C \):

\[
\text{Col}(C) = \bigsqcup_{C' \subset C} \text{Col}^*(C').
\]

This implies that

\[
Z_C = \sum_{C' \subset C} \left( \sum_{\sigma' \in \text{Col}^*(C')} |\sigma'|_C^{-1} A_C(\sigma') \right).
\]

Second, check that we have

\[
|\sigma'|_C = |\sigma'|_{C'} N_{C', C},
\]

with \( N_{C', C} \) the number of subfoams of \( C \) isomorphic to \( C' \). Third, use cylindrical consistency to obtain

\[
Z_C = \sum_{C' \subset C} N_{C', C}^{-1} Z_{C'}^\star
\]

and conclude to (2).

The LQG spinfoam amplitudes \( Z_C \) discussed in the literature do not satisfy the condition above. However, a simple modification of these amplitudes does. To see this, let us start from the Lorentzian version of the LQG amplitude defined in [7] and [8]. Using the representation given in [6], the amplitude of a foam with a given coloring \( \sigma = (j_f) \) is

\[
Z_C(\sigma) = \int_X dG \prod_f (2j_f + 1)^2 \chi^{j_f-j_f} (H_f(G)) \prod_{e \in \partial f} \chi^j (h_{j_f}(G)),
\]

where \( X \) is the group \( SL(2, \mathbb{C})^{n_C} \times SU(2)^{m_C} \) with \( n_C \) and \( m_C \) the two integers depending on \( C \), \( dG \) its Haar measure, \( H_f : X \rightarrow SL(2, \mathbb{C}), h_{j_f} : X \rightarrow SU(2) \) and \( \chi^j \) (resp. \( \chi^0 \)) the characters of \( SL(2, \mathbb{C}) \) (resp. \( SU(2) \)). See [5, 6] for the rest of the definition. It is immediate to see that if \( j_f = 0 \), then the face \( f \) gives no contribution to the amplitude, since \( (2j_f + 1) = 1 \) and \( \chi^{0,0} = \chi^0 = 1 \). Hence, \( Z_C(\sigma) \) is invariant under trivial extension. But the condition considered in the previous section was the invariance of (4), not the invariance of \( Z_C(\sigma) \). Thus, the spinfoam model (9) does not satisfy the condition for (3) to hold.

However, a slight modification of the model satisfies this condition. It suffices to replace (9) by

\[
Z_C(\sigma) \rightarrow |\sigma|_C^{-1} Z_C(\sigma).
\]

This modification does not interfere with the interesting physical properties of the model; for instance, it does not affect any of the results on the asymptotic limit of the theory [17–20]. But is there any rationale for considering such a modification and expecting it to be part of the quantum theory of gravity? There is. The spinfoam sum is meant to be an implementation of the Misner–Hawking sum over geometries,

\[
Z = \int_{\text{Metrics}/\text{Diff}} Dg_{\mu\nu} e^{\int_{\partial C}[R_{\mu\nu}]}.
\]

Here the integral is not over metrics, but over equivalence classes of metrics under diffeomorphisms. In the truncation induced by the choice of a foam, the diffeomorphisms are reduced to the automorphisms \( \phi \) of the foam. Therefore, the colorings \( \sigma \) and \( \sigma \circ \phi \) have a natural interpretation as the discrete residual equivalent of diffeomorphism-related metrics. This interpretation is reinforced by the fact that the amplitude is in fact invariant. If we want
to integrate over geometries, then the contributions of $\sigma$ and $\sigma \circ \phi$ represent an overcounting, and we must divide it by their number, namely by $|\sigma|_C$. The same conclusion can be reached by interpreting colored spinfoams as histories of non-trivial spin networks. Then colorings related by foam automorphisms clearly represent the same history. For both these reasons, it is physically interesting to consider the amplitudes modified as in (10). Then the sum over foams is equal to the continuum limit.

To be sure, if the multiplicities are not included, a relation between the models $Z$ and $Z^*$ still holds. This is not (2), but

$$Z_C = \sum_{C \leq C'} N_{C,C} Z_{C'}^*.$$  \hspace{1cm} (12)

That is, refining foams is still closely related to summing over foams, but less neatly. Observe also that the sum on the right-hand side is much less likely to converge in this case, as the number $N_{C,C'}$ grows very quickly with the number of vertices, edges and faces of $C$ and $C'$.

In conclusion, we observe that the convergence between the perturbative-QED and the lattice-QCD pictures should not be too surprising, for it follows from the physics of general relativity. Heuristically, lattice sites are small regions of space; according to general relativity, these are excitations of the gravitational field. Therefore they are themselves quanta of a (generally covariant) QFT. This scenario is realized concretely in LQG: an $N$-quanta state of gravity has the very same structure as a Yang–Mills state on a lattice with $N$ sites [6].

From this perspective, a foam is nothing but a ‘history’ of space quanta, in the sense in which a Feynman graph is a ‘history’ of field quanta. Technically, the convergence is realized by the fact that the faces of the foam carrying vanishing spin have no effect on the amplitude: a region where the gravitational field vanishes is like a non-existing region. This is different from lattice QCD, where a region of the lattice where the field vanishes cannot be eliminated from the lattice without altering the associated amplitude. In the latter case, if the field vanishes, what remains is empty space; in the former, if the field vanishes, what remains is nothing at all.

To summarize, we have given a precise definition of the continuum limit of a spinfoam model. We have observed that under certain general conditions, if this limit exists, it can equally be expressed as the sum over foams, by simply restricting the amplitudes to those with non-trivial spins. The condition for this to happen rules out a certain technical modification of the LQG spinfoam amplitude, and suggests to modify the spinfoam amplitude by factoring away the action of the automorphisms of the foam. These can be viewed as the residual action of the diffeomorphism group on the truncated theory. A question we have left open is the relationship between the combinatorial coefficients introduced by this factorization and those derived from the group field theory formulation of the theory [21–23].

Acknowledgments

We thank E Magliaro and C Perini for sending us a draft of their work, A Baratin for useful comments, J Lewandowski for pointing out an error in the first version of this draft and S Speziale for helping us correct it.

References

[1] Ambjorn J, Jurkiewicz J and Loll R 2010 Causal dynamical triangulations and the quest for quantum gravity arXiv:1004.0352 [hep-th]
[2] Regge T 1961 General relativity without coordinates Nuovo Cimento 19 558–71
[3] Ashtekar A, Rovelli C and Smolin L 1992 Weaving a classical geometry with quantum threads Phys. Rev. Lett. 69 237–40 (arXiv:hep-th/9203079)
[4] Rovelli C and Smolin L 1995 Spin networks and quantum gravity Phys. Rev. D 52 5743–59 (arXiv:gr-qc/9505006)
[5] Rovelli C 2011 Simple model for quantum general relativity from loop quantum gravity J. Phys. Conf. Ser. 314 012006 (arXiv:1010.1939)
[6] Rovelli C 2011 Zakopane lectures on loop gravity arXiv:1102.3660
[7] Engle J, Livine E, Pereira R and Rovelli C 2008 LQG vertex with finite Immirzi parameter Nucl. Phys. B 799 136–49 (arXiv:0711.0146)
[8] Freidel L and Krasnov K 2008 A new spin foam model for 4D gravity Class. Quantum Grav. 25 125018 (arXiv:0708.1595)
[9] Ding Y and Rovelli C 2010 Physical boundary Hilbert space and volume operator in the Lorentzian new spin-foam theory Class. Quantum Grav. 27 205003 (arXiv:1006.1294)
[10] Magliaro E and Perini C 2010 Local spin foams arXiv:1010.5227
[11] Bahr B, Hellmann F, Kaminski W, Kisielowski M and Lewandowski J 2011 Operator spin foam models Class. Quantum Grav. 28 105003 (arXiv:1010.4787 [gr-qc])
[12] Atiyah M 1988 Topological quantum field theory Publ. Math. IHES 68 175–86
[13] Atiyah M 1990 The Geometry and Physics of Knots (Cambridge: Cambridge University Press)
[14] Baez J C 1998 Spin foam models Class. Quantum Grav. 15 1827–58 (arXiv:gr-qc/9709052)
[15] Baez J C 2000 An introduction to spin foam models of BF theory and quantum gravity Lect. Notes Phys. 543 25–94 (arXiv:gr-qc/9905087)
[16] Oeckl R 2008 General boundary quantum field theory: foundations and probability interpretation Adv. Theor. Math. Phys. 12 319–52 (arXiv:hep-th/0509122)
[17] Barrett J W, Dowdall R J, Fairbairn W J, Gomes H and Hellmann F 2009 A summary of the asymptotic analysis for the EPRL amplitude arXiv:0909.1882
[18] Conrady F and Freidel L 2008 Path integral representation of spin foam models of 4D gravity Class. Quantum Grav. 25 245010 (arXiv:0806.4640)
[19] Bianchi E, Magliaro E and Perini C 2009 LQG propagator from the new spin foams Nucl. Phys. B 822 245–69 (arXiv:0905.4082)
[20] Rovelli C and Zhang M 2011 Euclidean three-point function in loop and perturbative gravity Class. Quantum Grav. 28 175010
[21] Geloun J B, Gurau R and Rivasseau V 2010 EPRL/FK group field theory Europhys. Lett. 92 60008 (arXiv:1008.0354)
[22] Krajewski T, Magnen J, Rivasseau V, Tanasa A and Vitale P 2011 Quantum corrections in the group field theory formulation of the EPRL/FK models Phys. Rev. D 83 104051 (arXiv:1007.3150)
[23] Baratin A, Girelli F and Oriti D 2011 Diffeomorphisms in group field theories arXiv:1101.0590