''Sum over surfaces'' form of loop quantum gravity

Michael P. Reisenberger
Instituto de Física, Universidad de la República, Tristan Narvaja 1674, 11200 Montevideo, Uruguay
and Erwin Schrödinger Institute, A-1090 Vienna, Austria

Carlo Rovelli*
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260;
Center for Gravity and Geometry, Penn State University, State College, Pennsylvania 16801;
and Erwin Schrödinger Institute, A-1090 Vienna, Austria
(Received 16 April 1997)

We derive a *spacetime* formulation of quantum general relativity from (Hamiltonian) loop quantum gravity. In particular, we study the quantum propagator that evolves the three-geometry in proper time. We show that the perturbation expansion of this operator is finite and computable order by order. By giving a graphical representation in the manner of Feynman of this expansion, we find that the theory can be expressed as a sum over topologically inequivalent (branched, colored) two-dimensional (2D) surfaces in 4D. The contribution of one surface to the sum is given by the product of one factor per branching point of the surface. Therefore branching points play the role of elementary vertices of the theory. Their value is determined by the matrix elements of the Hamiltonian constraint, which are known. The formulation we obtain can be viewed as a continuum version of Reisenberger's simplicial quantum gravity. Also, it has the same structure as the Ooguri-Crane-Yetter 4D topological field theory, with a few key differences that illuminate the relation between quantum gravity and topological quantum field theory. Finally, we suggest that certain new terms should be added to the Hamiltonian constraint in order to implement a ''crossing'' symmetry related to 4D diffeomorphism invariance. [S0556-2821(97)01718-9]

PACS number(s): 04.60.Ds, 04.60.Gw

I. INTRODUCTION

An old dream in quantum gravity [1] is to define a manifestly spacetime-covariant Feynman-style ''sum over trajectories'' [2], sufficiently well defined to yield finite results order by order in some expansion. The Hamiltonian theory has obtained encouraging successes in recent years, but it suffers for the well-known lack of transparency of the frozen time formalism, for the difficulty of writing physical observables, and for operator ordering ambiguities. These problems are related to the lack of manifest four-dimensional (4D) covariance. Here, we derive a covariant spacetime formalism *from* the Hamiltonian theory. This is of course the path followed by Feynman to introduce his sum over trajectories in the first place [3]. What we obtain is surprising: we obtain a formulation of quantum gravity as a sum over surfaces in spacetime. The surfaces capture the gravitational degrees of freedom. The formulation is ''topological'' in the sense that one must sum over topologically inequivalent surfaces only, and the contribution of each surface depends on its topology only.[1] This contribution is given by the product of elementary ''vertices,'' namely, points where the surface branches. The sum turns out to be finite and explicitly computable order by order. The main result of this paper is the construction of this finite ''sum over surfaces'' formulation of quantum gravity.

Let us sketch here the lines of the construction. Given gravitational data on a spacelike hypersurface Σ_i, the three-geometry on a surface Σ_f at a proper time T in the future of Σ_i (as measured along geodesics initially at rest on Σ_i) is uniquely determined in the classical theory. It is then natural to study the corresponding evolution operator $U(T)$, that propagates states from Σ_i to Σ_f in the quantum theory. This operator, first considered by Teitelboim [4], codes the dynamics of the quantum gravitational field, and is analogous to the Feynman-Nambu proper time propagator [5] for a relativistic particle. Here, we construct the operator $U(T)$ in quantum general relativity (GR), and we expand it in powers of T. We obtain a remarkable result: the expansion is finite order by order. This is our first result.

Next, we construct a graphical representation of the expansion. This is obtained by observing that topologically inequivalent colored 2D surfaces σ in spacetime provide a natural bookkeeping device for the terms of the expansion. We obtain an expression for $U(T)$ as a sum of terms labeled by surfaces σ bounded by initial and final states. A surface σ consists of simple components (two-manifolds), or ''faces,'' that carry a positive integer or color. Faces meet on 1D ''edges,'' colored as well. Edges, in turn, meet at branching points, denoted ''vertices.'' The weight of each surface in the sum is a product of factors associated to its vertices. The value of a vertex is determined by the Hamiltonian constraint, and is given by a simple function (involving Wigner $3n$-j symbols) of the colors of the adjacent faces and edges. This ''sum over surfaces'' version of the dynamics of quantum general relativity is our second result.

The construction allows us to consider transformation

*Electronic address: rovelli@pitt.edu

[1]More precisely: on the diffeomorphism-invariant properties of the surface.

properties of the Hamiltonian constraint under 4D diffeomorphisms (diff) in a manifestly covariant way. This analysis suggests the addition of certain new terms to the constraint, corresponding to an alternative operator ordering, which implements a ''crossing'' symmetry at the vertices, with a nice geometrical appeal. Thus, 4D diff invariance may be a key for reducing the present ambiguity in the operator ordering. Furthermore, the new terms seem to prevent some potential problems with locality pointed out by Smolin [6]. The introduction of these new terms in the quantum Hamiltonian is our third result.

The idea that one could express the dynamics of quantum gravity in terms of a sum over surfaces has been advocated in the past, particularly by Baez [7] and Reisenberger [8]. On the lattice, a sum over surfaces was recently developed by Reisenberger [9], and the lattice construction has guided us for the continuum case studied here. It is important to emphasize, however, that the present construction is entirely derived from the canonical quantum theory in the continuum.

The sum over surfaces we obtain has striking similarities with topological quantum field theory (TQFT). More precisely, it has the same kinematic as the Ooguri-Crane-Yetter model [10,11], a four-dimensional (4D) TQFT which extends the Ponzano-Regge-Turaev-Viro 3D TQFT [12–14] to four dimensions. Essentially, the difference is given just by the weight of the vertices. In Appendix B we discuss similarities and differences between the two theories. The discussion, we believe, sheds much light over the tantalizing issue [15] of the relation between finite-number-of-degrees-of-freedom TQFT and quantum gravity. In particular, we argue that a diff-invariant quantum field theory with an *infinite* number of local (but nonlocalized) degrees of freedom—such as quantum general relativity—can be obtained by having a sum over arbitrarily fine triangulations, instead of triangulation independence, as in combinatorial TQFT's.

On the other hand, the sum over surfaces we obtain can be viewed as (a first step towards) a concrete implementation of Hawking's sum over four-geometries [2]. In fact, the surfaces over which we sum have an immediate interpretation as ''quantum'' four-geometries, as we will illustrate. This fact should make the the general techniques of covariant generalized quantum mechanics [16] available to quantum gravity, potentially simplifying the difficulties with physical observables of the Hamiltonian formalism.

The basis of our construction is loop quantum gravity [17]. The finiteness of the sum-over-surfaces and the picture of a ''discrete four-geometry'' that emerges from this work are related to the fact that geometrical operators have discrete spectra. The discreteness of the spectra of area and volume—and the ''quantized'' structure of space that these spectra suggests—is a central result in loop quantum gravity, first obtained by Rovelli and Smolin in [18], and later confirmed and clarified by a number of authors [19–23]. The main ingredient of our construction is the quantum Hamiltonian constraint [17,24]. In particular, Thiemann's version of the Hamiltonian constraint [25] and some variants of it play an essential role here. Matrix elements of this operator have been computed explicitly in [26], using the methods developed in [21].

In Sec. II, we summarize the basics of nonperturbative loop quantum gravity. In Sec. III we define the proper time propagator and its expansion. In Sec. IV we show that the proper time propagator can be expressed as a sum over surfaces. In Sec. V we discuss crossing symmetry and the new terms of the Hamiltonian constraint. In Sec. VI we summarize and comment our results. Appendix A is a brief glossary of some geometrical terms employed. Appendix B contains the comparison with TQFT. In Appendix C we give an example of a 3D diff-invariant scalar product.

II. CANONICAL LOOP QUANTUM GRAVITY

A. Kinematic

We start with nonperturbative canonical quantum gravity in the loop representation [17]. The Hilbert space $\mathcal{H}$ of the theory is spanned by the basis $|S\rangle$, where S is a spin network [27]. A spin network is a colored graph Γ embedded in a (fixed) three-dimensional compact manifold Σ.[2] For a fixed choice of a treelike expansion at the nodes, these states are orthonormal [21,28,29] (see [21] for details on their normalization):

$$\langle S'|S\rangle = \delta_{SS'}; \tag{1}$$

the matrix elements of the change of basis between different nodes' expansion can be derived from Eq. (C7). An equivalent construction of this Hilbert space can be obtained in terms of functions over (generalized) connections [30]. For details on the equivalence between the two formalisms, see [31].

The dynamics of quantum general relativity is governed by two operators in $\mathcal{H}$: the Diff constraint operator $C[\vec{N}]$ and the Hamiltonian constraint $C_L[N]$. Let us examine them.

B. The diff constraint and its solutions

For every diffeomorphism $f:\Sigma\to\Sigma$ (in diff_0, the component of the diffeomorphism group connected to the identity), let $D[f]$ be the operator in $\mathcal{H}$ giving the natural action $f:S\mapsto f\cdot S$ of the diffeomorphism on the spin network states. Namely,

$$D[f]|S\rangle = |f\cdot S\rangle. \tag{2}$$

For every vector field $\vec{N}$ on Σ that generates a one parameter

[2]We recall the definition of coloring of a spin network [21]. Each node of the graph with valence higher than 3 (more than three adjacent links) is arbitrarily expanded in a treelike trivalent subgraph. The internal links of the subgraph are denoted virtual links. The coloring of the graph is an assignment of a positive integer to each real or virtual link—in such a way that at every trivalent node the sum of the three colors is even and none of the colors is larger than the sum of the other two. The set of the colorings of the virtual links of a node is also called coloring of the node. A coloring can be thought as an assignment of an irreducible SU(2) representation to each link and of an invariant coupling tensor to each node.

family f_t of diffeomorphisms (by $df_t/dt=\vec{N}$, f_0=identity)[3] the Diff constraint is defined by[4]

$$C[\vec{N}]=-i\frac{d}{dt}D[f_t]\Big|_{t=0} \tag{3}$$

and corresponds to the classical diffeomorphism constraint smeared with Shift function $\vec{N}$. The space $\mathcal{H}_{\rm diff}$ of the solutions of the diffeomorphism constraints is defined as $\mathcal{H}_{\rm diff}=\mathcal{H}/{\rm diff}_0$. It is spanned by a basis $|s\rangle$, where s is an s knot, namely an equivalence class of spin networks under diffeomorphisms, which define the linear structure of $\mathcal{H}_{\rm diff}$. One can define the scalar product in $\mathcal{H}_{\rm diff}$ by an integration [32,33] over ${\rm diff}_0$. If $S\in s$ and $S'\in s'$,

$$\langle s|s'\rangle=\mathcal{N}\int_{{\rm diff}_0}[df]\langle f\cdot S|S'\rangle. \tag{4}$$

$\mathcal{N}$ is a normalization factor. Equation (4) is meaningful because the integrand vanishes over most of the integration space (because two spin network states are orthogonal unless they have the same graph) and is constant on a discrete number of regions whose volume is normalized to one by $\mathcal{N}$. Thus we have

$$\langle s|s'\rangle=\sum_{\rho}\langle \rho S_i|S'\rangle, \tag{5}$$

where the sum is over the (discrete) automorphisms ρ that send the graph and the links' coloring into themselves. See Appendix C for an example, and Refs. [32,30] for a rigorous construction. It is useful to view an s-knot state as a group integral of a spin network state:

$$\langle s|=\mathcal{N}\int_{{\rm diff}_0}[df]\langle f\cdot S|, \tag{6}$$

where $S\in s$.

C. The Hamiltonian constraint

The Hamiltonian constraint that we consider is the density-weight 1 Hamiltonian density, smeared with a densit-weight 0 Lapse function N. The Lorentzian Hamiltonian constraint $C_L[N]$ can be written as the sum of two terms: $C_L[N]=C[N]+V[N]$ [34], where $C_L[N]$ is the Euclidean Hamiltonian constraint. For simplicity we deal here only with the first term. Thus, we are dealing below with Euclidean quantum gravity only. We expect the methods developed here to be extendible to the $V[N]$ term as well, and therefore to Lorentzian GR, using the techniques developed by Thiemann [25].

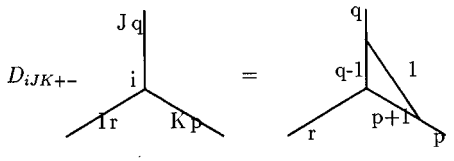


FIG. 1. Action of $D_{iJK\epsilon\epsilon'}$. r, q, and p are the colors of the links I, J, and K.

The definition of $C[N]$ is plagued by ordering ambiguities [24,25,35]. Some of these are fixed by 3D diff invariance [24]. In Sec. V we discuss how 4D diff invariance might fix others. Here, we recall Thiemann's version of the Hamiltonian constraint, which is the starting point of our construction. First, the nonsymmetric operator $C_{\rm ns}[N]$ is defined as

$$C_{\rm ns}[N]|S\rangle=2\sum_{i\in n(S)}N(x_i)\times\sum_{(JK)\in e(i)}\sum_{\epsilon=\pm1,\epsilon'=\pm1}A_{iJK\epsilon\epsilon'}(S)D_{iJK\epsilon\epsilon'}|S\rangle. \tag{7}$$

Here i labels the nodes of S [which form the set $n(S)$], x_i are the coordinates of the node i, (JK) labels the couples of distinct links adjacent to the node i [these form the set $e(i)$], and the operator $D_{iJK\epsilon\epsilon'}$ was introduced in [35]; it acts on the spin network by creating two new trivalent nodes i' and i'' on the the two links J and K, respectively, connected by a link with color 1, and adds ϵ (respectively ϵ') to the color of the link connecting i and i' (respectively i and i''). This is illustrated in Fig. 1.

The precise location of the nodes and the link added is an arbitrary regularization choice. The coefficients $A_{iIJ\epsilon\epsilon'}(S)$ of Thiemann's operator are well defined and can be computed explicitly [26]. They are functions of the colors of the links adjacent Tothe node i; they are finite and can be expressed as products of linear combinations of n-j symbols of SU(2).

It is important to notice that Thiemann's operator was derived using the infinite-dimensional differential-geometry techniques introduced in [30]. These differential techniques were introduced as a mathematical systematization of the ideas on loop quantization introduced in [17]. They have shed much light on loop quantum gravity, have provided a rigorous mathematical foundation of the theory, and have led, among other results, to Thiemann's operator. However, the operator itself is a well-defined *algebraic* operator on the spin network basis, and the computation of its matrix elements is easier using algebra than using infinite-dimensional differential geometry [21]. The equivalence between the purely algebraic formalism (''loop representation'' or ''spin network representation'') and the differential formalism (''connection representation'') is shown in detail in [31] (see also [29,23]). The situation is analogous to the two well-known ways of computing the spectrum of the harmonic oscillator: one can use Dirac's *algebraic* technique, in the

[3]We put an arrow over vectors ($\vec{N}$), but not over spatial coordinates (x) or diffeomorphisms (f).

[4]Rigorously speaking, $C[\vec{N}]$ is not well defined on $\mathcal{H}$. This is due to funny (kinematical) inner product (1), in terms of which the action of the diffeomorphism group is not strongly continuous. This fact does not disturb the construction of the theory, because the only role played by $C[\vec{N}]$ is to implement invariance under the finite transformations it generates. These are well defined [30]. Here, it is useful to consider $C[\vec{N}]$ as well, because it plays a role in the formal manipulations below.

$|n\rangle$ basis or, alternatively, one can solve the *differential* Schrödinger equation, in the coordinate basis. Each basis has its own advantages, but the two formalisms are equivalent, and there is no sense in which one representation is more ''rigorous'' than the other. In particular, spectra of area and volume in quantum gravity can be computed either using algebraic techniques (this is the way they were *first* computed and their discreteness was discovered in [18]) or using differential techniques (see [23], and references therein). The resulting spectra are, of course, equal [22].

The full Euclidean Hamiltonian constraint $C[N]$ is then defined by symmetrizing $C_{\rm ns}[N]$:[5]

$$\langle S'|C[N]|S\rangle=\frac{1}{2}(\langle S'|C_{\rm ns}[N]|S\rangle+\langle S|C_{\rm ns}|S'\rangle). \quad (8)$$

Explicitly, we have

$$\begin{aligned}\langle S'|C[N]|S\rangle=&\sum_{i\in n(S)}N(x_i)\sum_{(JK)\epsilon\epsilon'}A_{iJK\epsilon\epsilon'}(S)\\&\times\langle S'|D_{iJK\epsilon\epsilon'}|S\rangle\\&+\sum_{i'\in n(S')}N(x_{i'})\sum_{(JK)\epsilon\epsilon'}A_{i'JK\epsilon,\epsilon'}(S)\\&\times\langle S'|D^{\dagger}_{i'JK\epsilon\epsilon'}|S\rangle, \quad (9)\end{aligned}$$

where

$$\langle S'|D^{\dagger}_{iJK\epsilon\epsilon'}|S\rangle=\langle S|D_{iJK\epsilon\epsilon'}|S'\rangle. \quad (10)$$

Notice that the Hamiltonian constraint is ''local,'' in the following sense. Given a spin network S, we may cut it in two parts, by cutting n links, obtaining two spin networks with open ends $\widetilde{S}$ and $\hat{S}$. Imagine we have two spin networks S_i and S_f that can be cut as $\widetilde{S}_i$ and $\hat{S}_i$, and, respectively, $\widetilde{S}_f$ and $\hat{S}_f$. Imagine that $\hat{S}_i=\hat{S}_f$. Then the matrix elements $\langle S_f|C[N]|S_i\rangle$ do not depend on the ''hat'' components $\hat{S}_i$ and $\hat{S}_f$, so we can write

$$\langle S_f|C[N]|S_i\rangle=\langle\widetilde{S}_f|C[N]|\widetilde{S}_i\rangle. \quad (11)$$

This decomposition will play a role below.

We simplify notation by introducing a single discrete index $\alpha,\beta\ldots$, to replace the discrete set of indices $(i,JK,\epsilon,\epsilon')$. For every spin network S, α ranges over a finite set $[S]$ of values, with

$$4\prod_{i\in n(S)}\frac{v_i(v_i-1)}{2} \quad (12)$$

values, where v_i is the valence of the node i. We also indicate by x_α the coordinates of the node with index α. Using this, we have

$$\begin{aligned}\langle S'|C[N]|S\rangle=&\sum_{\alpha\in[S]}N(x_\alpha)A_\alpha(S)\langle S'|D_\alpha|S\rangle\\&+\sum_{\beta\in[S']}N(x_\beta)A_\beta(S)\langle S'|D^{\dagger}_\beta|S\rangle. \quad (13)\end{aligned}$$

The Hamiltonian constraint transforms covariantly under the diffeomorphisms generated by the diff constraint

$$\{C[\vec{N}],C[N]\}=C[\mathcal{L}_{\vec{N}}N], \quad (14)$$

where $\mathcal{L}_{\vec{N}}$ is the Lie derivative along $\vec{N}$. Under a finite diffeomorphism f, we have

$$D[f]C[N]D^{-1}[f]=C[N_f], \quad (15)$$

where N_f is the transformed Lapse:

$$N_f(x)\equiv N[f(x)]. \quad (16)$$

The transformation properties of $C[N]$ under 4D diffeomorphisms are less clear. In the canonical formalism these are controlled by the commutator of $C[N]$ with itself, which, however, is not fully under control, due to the interplay between regularization and 3D diff invariance [36]. (Notice that in [25] it is shown that the commutator $\{C_{\rm ns}[N],C_{\rm ns}[M]\}$ vanishes on diff-invariant states; the commutator $\{C[N],C[M]\}$ is more tricky.) In Sec. IV we suggest a way for addressing these difficulties.

III. PROPER TIME EVOLUTION OPERATOR $U(T)$

Consider, as an illustrative example, the Schrödinger equation for a single particle in a potential. If H is the Hamiltonian operator, the equation is formally solved by the evolution operator

$$U(t)=U(t,0)=e^{-i\int_0^t dt'H(t')}, \quad (17)$$

where exponentiation, here and below, is time ordered. The matrix elements of this operator between position eigenstates define the propagator

$$P(\vec{x},t;\vec{x}',t')=\langle\vec{x}|U(t,t')|\vec{x}\rangle, \quad (18)$$

from which the solution of the Schrödinger equation with initial data $\psi(\vec{x}',t')$ at t' can be obtained by simple integration:

$$\psi(\vec{x},t)=\int dx'P(\vec{x},t;\vec{x}'t')\psi(\vec{x}',t'). \quad (19)$$

Under suitable conditions, the propagator can be computed by means of a perturbation expansion in the potential, and the expansion has a nice graphical representation.

For a (free) relativistic particle, we have the option between using the above formalism with the relativistic Hamiltonian ($H=\sqrt{\vec{p}^2+m^2}$), or using a manifestly covariant formalism. This was originally done by Feynman by changing the description of the dynamics: instead of representing motion by means of the evolution of the *three* variables $\vec{x}$ in t, we consider a (''fictitious'') evolution of the *four* variables

[5]Here we focus on Thiemann's *symmetric* operator.

$x=(\vec{x},t)$ in the proper time T. This evolution is generated by the operator $H=p^2-m^2=(p^0)^2-\vec{p}^2-m^2$. The corresponding proper time evolution operator and proper time propagator are

$$U(T)=e^{-i\int_0^T dt' H(t')} \tag{20}$$

and

$$P(\vec{x},t;\vec{x}',t';T)=\langle \vec{x},t|U(T)|\vec{x}',t'\rangle. \tag{21}$$

The relation between this proper time propagator and the physical propagator (which is the quantity we compare experiments with) is given by

$$P(\vec{x},t;\vec{x}',t')=\int_0^\infty dT\, P(\vec{x},t;\vec{x}',t';T) \tag{22}$$

or

$$P(\vec{x},t;\vec{x}',t')=\langle \vec{x},t|U|\vec{x}',t'\rangle, \tag{23}$$

where

$$U=\int_0^\infty dT\, U(T). \tag{24}$$

This can be verified by means of a simple calculation. U is the projector on the physical state space, which codes the theory's dynamics.

Alternatively, one can consider evolution in a fully arbitrary parameter t. Such evolution is generated by $H(t)=N(t)H$, where $N(t)$ is an arbitrary Lapse function. The corresponding evolution operator is

$$U_N=\exp\left(-i\int_0^1 dt\, N(t)H\right), \tag{25}$$

which is related to the physical U by the functional integral

$$U=\int [dN]U_N. \tag{26}$$

This functional integration can be split into two parts by defining the proper time T in terms of the Lapse as

$$T=\int_0^1 N(t)dt. \tag{27}$$

Using this, we can first integrate U_N over all Lapses N having the same T:

$$\widetilde{U}(T)=\int_T dN\, U_N, \tag{28}$$

where the subscript T indicates that the functional integral must be performed over all N's satisfying Eq. (27). Then we integrate over T to get the physical quantity U. Thus we have

$$U_N\mapsto\widetilde{U}(T)\mapsto U. \tag{29}$$

An important observation is that in order to compute the functional integral (28) we can simply *gauge fix N*, requiring, for instance $dN(t)/dt=0$. In this gauge, the integral becomes trivial, and we have $\widetilde{U}(T)=U(T)$, which is given in Eq. (20). In fact, the functional integration (26) over N is largely trivial, since U_N depends on N only via T.

We are now going to follow the same path in general relativity. In particular, we will concentrate here on the definition and the computation of the proper time evolution operator $U(T)$ and the corresponding proper time propagator (its matrix elements) for quantum general relativity.

A. Definition and meaning of the proper time propagator in general relativity

In the canonical theory, the (''unphysical'' or coordinate) evolution of the gravitational field is generated by the Hamiltonian

$$H_{N,\vec{N}}(t)=\int_\Sigma d^3x[N(t,x)C(x)+N^a(t,x)C_a(x)]$$
$$\equiv C[N(t)]+C[\vec{N}(t)] \tag{30}$$

(units are fixed here by $\hbar=c=16\pi G_{\text{Newton}}=1$, and we take Σ compact). The quantum evolution operator that evolves from an initial hypersurface Σ_i at $t=0$ to a final hypersurface Σ_f at $t=1$ is

$$U_{N,\vec{N}}=\exp -i\int_0^1 dt\, H_{N\vec{N}}(t). \tag{31}$$

We define the proper time evolution operator for quantum gravity as

$$U(T)=\int_{T,*}[dN,d\vec{N}]U_{N,\vec{N}}, \tag{32}$$

where the subscript $\{T,*\}$ means that the integral is over all shifts and lapses that satisfy

$$N(x,t)=N(t), \tag{33}$$

$$\int_0^1 dt\, N(x,t)=T. \tag{34}$$

Notice that T is the proper time separation between Σ_i and Σ_f, defined as the reading on Σ_f of the free falling test clock that started off at rest on Σ_i. This is because if the lapse is constant the geodesics that define the proper time foliation remains normal to the ADM hypersurfaces. We denote the matrix elements of the operator (32),

$$P(s_f,s_i;T)=\langle s_f|U(T)|s_i\rangle, \tag{35}$$

as the proper time propagator [4]. In this paper, we focus on this quantity. We compute it as a power expansion in T in the next subsection, and show that it admits a sum over surfaces representation in Sec. IV.[6]

The construction generalizes to a multifingered proper time. In this case, let Σ_f be given by $t=t(x)$. The coordinate time evolution operator from Σ_i to Σ_f at fixed lapse and shift is

$$U_{N,\vec{N}}=\exp\left(-i\int_\Sigma d^3x\int_0^{t(x)}dt[N(t,x)C(x)+N^a(t,x)C_a(x)]\right). \tag{36}$$

The multifingered proper time evolution operator is

$$U[T]=\int_{[T]}[dN,d\vec{N}]U_{N,\vec{N}}, \tag{37}$$

where the subscript $[T]$ means that the integral is over all shifts and lapses that satisfy Eq. (33) and

$$\int_0^{t[g_t^{-1}(x)]}dt\,N[g_t^{-1}(x),t]=T(x), \tag{38}$$

where g_t is the finite, time-dependent transformation of spacial coordinates generated by integrating the shift:

$$g_0(x)=x,\quad \frac{dg_t(x)}{dt}=\vec{N}(x,t),\quad g\equiv g_1. \tag{39}$$

g_t and g are functionals of the shift. As before, $T(x)$ gives the proper time separation between the two hypersurfaces, defined as the proper distance of Σ_f from Σ_i along a geodesic starting at rest on Σ_i on x Indeed, if the lapse is spacially constant, the geodesic will be $g_t^{-1}(x)$ (because in the coordinates $(\vec{y},t)=[g_t(x),t]$ the Shift vanishes, the Lapse is still constant and therefore $\vec{y}=$ const is a geodesic normal to all ADM slices). And therefore the geodesic that starts off in x reaches Σ_f at the time t determined by $t=t[g_t^{-1}(x)]$. Notice that the two notations $U[T]$ and $U(T)$ indicate different objects. $U[T]$ is the *multifingered* proper time propagator, a functional of the function $T(x)$, while $U(T)$ is the proper time evolution operator, which is the value of $U[T]$ for $T(x)=\text{const}=T$.

In the rest of this subsection, we discuss the physical meaning of the quantity we have defined and its role in the theory. First of all, the proper time propagator is the first step toward the computation of the physical evolution operator U, in the same sense as the Feynamn-Nambu proper time propagator. The operator U is given by functionally integrating $U_{N,\vec{N}}$ over *all* lapses and shifts:

$$U=\int[dN]\int[d\vec{N}]U_{N,\vec{N}}. \tag{40}$$

This functional integrations corresponds to the implementation of the canonical constraints. As for the relativistic particle considered in the previous section, we can split the computation of U from $U_{N,\vec{N}}$ into two steps,

$$U_{N,\vec{N}}\mapsto\tilde{U}[T]\mapsto U, \tag{41}$$

by first computing the propagator $\tilde{U}[T]$ at fixed $T(x)$ and then integrating over $T(x)$. As for the particle, we can partially fix the gauge in which we compute $U[T]$. In particular, we can choose to integrate over spacially constant lapses only. Therefore we have $\tilde{U}[T]=U[T]$, which is given in Eq. (37). And, as for the particle, we can write

$$U=\int[dT]U[T]. \tag{42}$$

Thus, U is just the integral over proper time of the multifingered proper time evolution operator (37).

We add a general argument that better illustrates why we can fix a gauge for computing $\tilde{U}[T]$. This argument is formal, but it is interesting because it illuminates the relation between what we are doing and the sum over geometries considered by Hawking [2]. In the metric formulation of canonical GR, U can be written as a sum over four-metrics bounded by given initial and final three-geometries. Each such four-metric determines a proper time separation $T(x)$ between the initial and final hypersurfaces. Therefore, the integral can be split in two parts, first the integral $\tilde{U}(T)$, restricted to four-metrics with total multifingered elapsed time $T(x)$, then the integration over $T(x)$. In computing $U(T)$, we can change integration variables from the four-metric to the ADM variables, namely, three-metric, lapse and shift. The integral contains a high redundancy, corresponding to diffeomorphism gauge-invariance (as the corresponding integral for the particle in the previous section did). We can fix part of this redundancy with a condition on the lapse. If we pick an arbitrary lapse, the condition that the proper time separation of the initial and final slices is $T(x)$ becomes a condition on the three-metrics over which we are integrating. However, we can choose a spacially constant lapse, satisfying Eqs. (33) and (38). We can always do that, because we may always slice a four-geometry with equal proper time hypersurfaces, with the result that the corresponding lapse and shift satisfy Eqs. (33) and (38). Conversely, any history of three-metrics, together with a lapse and shift satisfying Eqs. (33) and (38) defines a four-geometry with elapsed

[6]Generally, slicing by equal proper time hypersurfaces develops singularities, because the geodesics that define the proper time intersect. This causes the canonical evolution to break down: a coordinate system based on the slicing develops coordinate singularities where the Arnowitt-Deser-Misner (ADM) momentum density diverges. We ignore these difficulties here, but two comments are in order. First, the explicit expression for $U(T)$ that we obtain is simple and well defined order by order for any T. We think that potential singularities in $U(T)$ should be looked for directly in the quantumformalism. Second, Lewandowski [37] has pointed out that the classical evolution of the Ashtekar's variables is better behaved than the ADM variables, because Ashtekar's variables can be represented as differential forms, whose components are well behaved at coordinate sigularities of the type being considered.

proper time $T(x)$. In this way, we implement the $T(x)$ condition, without interfering with the integration over three-metrics. Thus, $\widetilde{U}[T]$ can be computed by fixing a laspe satisfying Eqs. (33) and (38).

The operator U is a key quantity for the theory. Computing it virtually amounts to solving the quantum constraints, including the Hamiltonian constraint, which codes all the dynamics of the theory. There are various of ways of looking at U. First of all, it is the projector on the the physical state space of the theory. Second, the scalar product $\langle s|U|s'\rangle$ defines the *physical* scalar product of the theory. Therefore,

$$\langle s,s'\rangle_{\text{physical}}=\int [dT]\langle s|U[T]|s'\rangle. \tag{43}$$

Finally, we can view matrix elements of U as observable transition amplitudes between quantum states. The details of the interpretaton of U will be discussed elsewhere, but in all these instances, the role of U is just analogous to its counterpart for the single particle. (In this paper we do not attempt to compute U.)

Does the proper time propagator $P(s_f,s_i;T)$, have a direct physical interpretation? A simple answer is that $P(s_f,s_i;T)$ codes the dynamics of the theory, but it has no direct physical meaning: only after integration over proper time we obtain a quantity that we can, in principle, compare with experiments.

This said, we *can* nevertheless assign a plausible physical interpretation to the proper time propagator $P(s_f,s_i;T)$, with caution. This would be particularly useful for helping intuition. Let us return to the relativistic particle. In that case, Feynman considered the ''fictitious'' evolution of x and t in the proper time T. Classically, this is not incorrect [because the equations of motion of the (4+1)-dimensional theory give the correct physical 3+1 motion] provided that one remembers that the degrees of freedom are 3 and not 4. Quantum mechanically, in the fictitious theory we are quantizing one variable too much. Taken literarly, the particle proper time propagator describes the three degrees of freedom of the particle position, plus the degree of freedom of an extra quantum variable sitting on the particle, growing with proper time, and not affecting the particle's motion.[7] With such a little clock on the particle (say the particle is an oscillating molecule), we could make experiments we could compare the proper time propagator with.

In general relativity, such a ''fictitious'' evolution with extra degrees of freedom is provided by the so-called ''local interpretation'' of the theory (see [38] for a detailed discussion). In this interpretation, the coordinates are interpreted as labels of reference-system (RS) physical objects. It follows that local quantities are physical observables, and that the lack of determinism of the Einstein equations can be interpreted as a consequence of the fact that the dynamical equations of the RS objects are neglected. Under this interpretation, GR is approximate (because we disregard the RS objects energy momentum) and incomplete (because we disregard the RS objects dynamical equations). The incompleteness leads to the apparent physical indeterminism. If we adopt this view, then we can say that s, s', and T are observable, because Σ_i can be physically specified by the RS objects, and we can use RS clocks to find out where Σ_f is. In doing so, we take approximations that might be ungranted (on the quantum behavior of the RS objects). Concretely, one may consider a definite model, for instance, the ''dust'' model introduced in [38] and studied in [39]. In such a model, a Σ_i to Σ_f propagator (where Σ_i is defined as dust variables) is an observable quantity. In a suitable limit in which the dust physical effects are diregarded, such a propagator might be approximated by the pure gravity proper time propagator $P(s_f,s_i;T)$. With all these caveats, one can intuitively think of $P(s_f,s_i;T)$ as the quantum amplitude that the quantum gravitational field be in the state s_f, T seconds after being in the state s_i.

B. Expansion of the proper time propagator

We begin with an observation. Let us write $U_{N,\vec{N}}$ as a limit of products of small time propagators. Writing $\epsilon=1/K$ and $t_k=k\epsilon$, for integers K and $k=1,\ldots,K$, we have

$$\begin{aligned}
U_{N,\vec{N}} &= \lim_{K\to\infty} e^{-i\epsilon H_{N,\vec{N}}(t_K)}\cdots e^{-i\epsilon H_{N,\vec{N}}(t_2)}e^{-i\epsilon H_{N,\vec{N}}(t_1)}\\
&= \lim_{K\to\infty} e^{-i\epsilon C[N(t_K)]}e^{-i\epsilon C[\vec{N}(t_K)]}\cdots e^{-i\epsilon C[N(t_2)]}e^{-i\epsilon C[\vec{N}(t_2)]}e^{-i\epsilon C[N(t_1)]}e^{-i\epsilon C[\vec{N}(t_1)]}\\
&= \lim_{K\to\infty} e^{-i\epsilon C[N(t_K)]}D[f_K]\cdots e^{-i\epsilon C[N_{t_2}]}D[f_2]e^{-i\epsilon C[N_{t_1}]}D[f_1]\\
&= \lim_{K\to\infty} e^{-i\epsilon C[N(t_K)]}D[f_K]\cdots\{(D[f_2]D[f_1])^{-1}e^{-i\epsilon C[N(t_2)]}D[f_2]D[f_1]\}(D^{-1}[f_1]e^{-i\epsilon C[N(t_1)]}D[f_1])\\
&= \lim_{K\to\infty} D[g](D^{-1}[g_K]e^{-i\epsilon C[N(t_K)]}D[g_K])\cdots(D^{-1}[g_2]e^{-i\epsilon C[N(t_1)]}D[g_2])(D^{-1}[g_1]e^{-i\epsilon C[N(t_1)]}D[g_1])\\
&= \lim_{K\to\infty} D[g]e^{-i\epsilon C[N_{\vec{N}}(t_K)]}\cdots e^{-i\epsilon C[N_{\vec{N}}(t_2)]}e^{-i\epsilon C[N_{\vec{N}}(t_1)]}=D[g]U_{N_{\vec{N}},0}.
\end{aligned} \tag{44}$$

[7] More precisely, the extra variable is not the evolution of this proper-time clock-variable T in a Lorentz time x^0, but rather the evolution of x^0 in T.

Here $N_{\vec{N}}$ is defined by $N(x,t)=N[g_t(x),t]$; namely, it is the lapse in the coordinates obtained by integrating the shift. f_t is the ''small'' diffeomorphism generated by the shift between the slices $t-1$ and t, while g_t is the finite diffeomorphism generated by the shift between the slices $t=0$ and t. The first equality in Eq. (44) is just one of the definitions of the time ordered exponential. The second is based on the fact that for sufficiently small time interval $1/K$ (sufficiently high K) one can disregard the commutator term in disentangling the exponent (this term is quadratic in $1/K$). The third equality is simply a rewriting of the exponent of an infinitesimal diffeomorphism as a finite (but ''small'') diffeomorphism. The fourth equality is simply the insertion of terms such as $(D[f_1]D^{-1}[f_1])$ in suitable places. The fifth equality is the replacement of sequences of spatial diffeomorphisms $(D[f_n]\cdots D[f_2]D[f_1])$ by their product, which is $(D[g_n])$. The penultimate equality is the key one; it follows directly from Eq. (15); namely, from the transformation properties of the Hamiltonian constraint under spatial diffeomorphisms. The last equality follows again from the definition of ordered exponential. In other words, we have shown that the temporal evolution generated by the lapse and the evolution generated by the shift can be disentangled.

While the manipulations above are formal (they are made inside a limit), the result itself is geometrically obvious: we can always rearrange the coordinates so that the shift is zero, and compensate with a finite change of space coordinates at the end. If we do so, the lapse N must be replaced by the lapse in the new coordinates, which is $N_{\vec{N}}$.

If the lapse is constant in space, $N_{\vec{N}}=N$. Then $U_{N,0}$ can be expanded as

$$U_{N0}=1+(-i)\int_0^\tau dt\, C[N(t)] + (-i)^2\int_0^\tau dt\int_t^\tau dt' C[N(t')]C[N(t)]+\cdots. \tag{45}$$

Its matrix elements between two spin network states can be expanded as

$$\langle S_f|U_{N,0}(T)|S_i\rangle=\langle S_f|S_i\rangle+(-i)\int_0^\tau dt\langle S_f|C[N(t)]|S_i\rangle + (-i)^2\int_0^\tau dt\int_t^\tau dt'\langle S_f|C[N(t')]|S_1\rangle \times\langle S_1|C[N(t)]|S_i\rangle+\cdots \tag{46}$$

where we have inserted a complete set of intermediate states $|S_1\rangle\langle S_1|$ (over which summation is understood). Using the explicit form (13) of the Hamiltonian constraint operator, we have

$$\langle S_f|U_{N,0}|S_i\rangle = \langle S_f|S_i\rangle+(-i)\int_0^\tau dt\Bigg(\sum_{\alpha\in[S_i]} N(t,x_\alpha)A_\alpha(S_i)\langle S_f|D_\alpha|S_i\rangle + \sum_{\beta\in[S_f]} N(t,x_\beta)A_\beta(S_f)\langle S_f|D_\beta^\dagger|S_i\rangle\Bigg) + (-i)^2\int_0^\tau dt\int_t^\tau dt'\sum_{\alpha\in[S_i]}\sum_{\alpha'\in[S_1]} \times N(t,x_\alpha)N(t',x_{\alpha'})A_\alpha(S_i)A_{\alpha'}(S_1)\langle S_f|D_\alpha|S_1\rangle \times\langle S_1|D_{\alpha'}|S_i\rangle+\cdots \tag{47}$$

(the second order term has three more summands, corresponding to the $DD^\dagger, D^\dagger D, D^\dagger D^\dagger$ terms). The first point to be noticed in this expression is that the sum over the intermediate state S_1 is finite. This is because both D and $D^\dagger$ yield a finite number of terms only, when acting on a spin network state.[8] Thus, the above expression is finite order by order. Next, the integrations can be performed explicitly, using Eq. (34). We obtain

$$\langle S_f|U|S_i\rangle=\langle S_f|S_i\rangle+(-iT)\Bigg(\sum_{\alpha\in[S_i]} A_\alpha(S_i)\langle S_f|D_\alpha|S_i\rangle + \sum_{\alpha\in[S_f]} A_\alpha(S_f)\langle S_f|D_\alpha^\dagger|S_i\rangle\Bigg) + \frac{(-iT)^2}{2!}\sum_{\alpha\in[S_i]}\sum_{\alpha'\in[S_1]} A_\alpha(S_i)A_{\alpha'}(S_1) \times\langle S_f|\hat{D}_{\alpha'}|S_1\rangle\langle S_1|\hat{D}_\alpha|S_i\rangle+\cdots. \tag{48}$$

The structure of the expansion is now rather clear. At each order n, we have the D operator acting n times, n factors A, and a *finite* number of terms, coming from summing over nodes, links and $\epsilon=\pm1$.

Our next step is to integrate over shift and lapse [satisfying Eq. (38)]. The integration over lapse is trivial, as its dependence has dropped out the integral. This confirms the independence from the laspe that was mentioned in the previous section. The integration over the shift amounts to imposing the diff constraint. Indeed, it turns out to be equivalent to an integration over the diffeomorphism group, as in the group integration technique for solving the diff constraint. Using Eq. (44), we have

$$U(T)=\mathcal{N}'\int[dN]\int[d\vec{N}]D[g[\vec{N}]]U_{N_{\vec{N}},0}(T), \tag{49}$$

where we have explicitly indicated the dependence of g on $\vec{N}$ for clarity. We change integration variable $N\to N_{\vec{N}}$ (the Jacobian must be 1, since this amounts to a change of coordinates), and obtain

[8] Lewandowski [40] noticed that this finiteness might fail because of the moduli parameters of high valent intersections which were studied in [41]. The role of these parameters in the theory, however, is unclear. Finiteness of the proper time expansion may indicate that the correct version of the theory is the one in which the moduli parameters are removed, as suggested by many, and recently detailed in [42].

$$U(T)=\mathcal{N}'\left(\int [dN]U_{N,0}(T)\right)\left(\int [d\vec{N}]D[g[\vec{N}]]\right). \tag{50}$$

The $\vec{N}$ integration can be traded for an integration over diff_0 changing variables from $\vec{N}$ to $g[\vec{N}]$, so we obtain

$$U(T)=\mathcal{N}\int_{\mathrm{diff}_0}[dg]D[g]U_{N,0}(T) \tag{51}$$

for an arbitrary (irrelevant) choice of N satisfying Eq. (34), say $N=T$. The matrix elements of this operator are given by

$$\langle S_f|U(T)|S_i\rangle=\mathcal{N}\int_{\mathrm{diff}_0}[dg]\langle g\cdot S_f|U_{N,0}(T)|S_i\rangle. \tag{52}$$

The operator $U(T)$ is now well defined in $\mathcal{H}_{\mathrm{diff}}$. Indeed, it is immediate to see that it is diff invariant. For every two s-knots s_i and s_f in $\mathcal{H}_{\mathrm{diff}}$, we can arbitrarily pick S_i and S_f such that $S_i\in s_i$ and $S_f\in s_f$, and we have the key result that

$$\langle s_f|U(T)|s_i\rangle\equiv\langle S_f|U(T)|S_i\rangle \tag{53}$$

is well defined (independent from the S_i and S_f chosen).

Furthermore, the operator D, depends on an arbitrary regularization—the location of the added link—but a moment of reflection shows that the dependence on the regularization drops out in the step from $U_{N,\vec{N}}(T)$ to $U(T)$, by integrating the shift. The reason is that different regularizations are related to each other by a finite diffeomorphism: the states $D_\alpha|S\rangle$ and $D'_\alpha|S\rangle$, where D and D' indicate two different regularizations of D are in the same s knot: their difference becomes irrelevant in the scalar product (53). This result is due to the fact that all the factors in the expansion are individually well defined at the diffeomorphism invariant level. More precisely we have that

$$\langle s_f|\sum_{\alpha\in[s_i]}A_\alpha(s_i)D_\alpha|s_i\rangle \tag{54}$$

is not only well defined, but also independent from the regularization of D. This fact allows us to write our expansion directly in diff-invariant form as

$$\begin{aligned}\langle s_f|U(T)|s_i\rangle=\langle s_f|s_i\rangle+(-iT)\Bigg(&\sum_{\alpha\in[s_i]}A_\alpha(s_i)\langle s_f|D_\alpha|s_i\rangle\\&+\sum_{\alpha\in[s_f]}A_\alpha(s_f)\langle s_f|D^\dagger_\alpha|s_i\rangle\Bigg)\\&+\frac{(-iT)^2}{2!}\sum_{\alpha\in[s_i]}\sum_{\alpha'\in[s_1]}A_\alpha(s_i)A_{\alpha'}(s_1)\\&\times\langle s_f|D_{\alpha'}|s_1\rangle\langle s_1|\hat{D}_\alpha|s_i\rangle+\cdots.\end{aligned} \tag{55}$$

This expression gives the three-geometry to three-geometry propagator of quantum general relativity as a series finite at every order. Notice that the expansion is in power of T or, equivalently, in inverse powers of the Planck length, because this must divide T in order to recover physical dimensions. The utility of a perturbation expansion in inverse powers of G has been advocated in quantum gravity by Isham, Teitelboim, and others [43]. Intuitively, we can think that this quantity represents the probability amplitude that if we have a quantum state of the gravitational field (a quantum three-geometry) $|s_i\rangle$ over a surface Σ_i, we will find the quantum three-geometry $|s_f\rangle$ on the surface Σ_f in a proper time T.

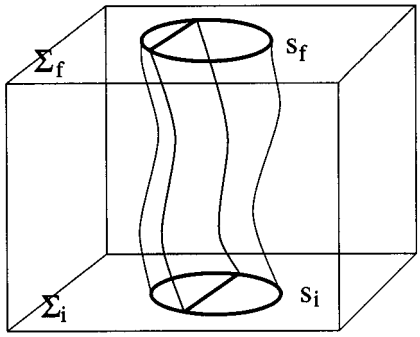


FIG. 2. Surface corresponding to a term of order zero.

IV. SUM OVER SURFACES

Surfaces in spacetime provide a natural bookkeeping device for the terms of the expansion (55) in the same manner in which Feynman graphs provide a bookkeeping device for conventional QFT perturbation expansion. This fact leads us to give a nice graphical interpretation to the expansion (55).

Consider the 4D manifold $\mathcal{M}=[0,1]\times\Sigma$. Denote the two connected components of the boundary of $\mathcal{M}$ as Σ_i and Σ_f. We now associate a 2D colored surface σ in $\mathcal{M}$—defined up to 4D diffeomorphisms—to each nonvanishing term of the sum in the right-hand side of Eq. (55). We begin by drawing s_i in Σ_i and s_f in Σ_f. As s_i and s_f have no information about the actual location of the graph, location is chosen arbitrarily (that is, up to a diffeomorphism in diff_0). The first term in Eq. (55) (zeroth order in T) is nonvanishing only if $s_f=s_i$. In this case, let us slide s_i across $\mathcal{M}$ from Σ_i to Σ_f, in such a way that it ends up over s_f. To the term of order zero we associate the surface $\sigma=s_i\times[0,1]$ swept by s_i, see Fig. 2.

Notice that this is possible because the two spin networks are in the same s knot. The surface we obtain is formed by 2D faces—submanifolds of $\mathcal{M}$—joined along edges. The faces are swept by the spin network links, and the edges are swept by the spin network nodes. We color every face with the color of the corresponding link of s_i, and every edge with the color of the corresponding node of s_i.

The surface associated to one of the summands of the second term (first order in T) in Eq. (55) is then defined as follows. In each summand, one of the nodes of s_i, say the node i, is altered by the operator D. s_f has two nodes more than s_i, say i' and i''. We begin by sliding s_i into the manifold by an arbitrary finite amount, until a position, say s. Let p be the point in which the node i ends up. Then we slide s_f from Σ_f through the manifold in such a way that it converges to s. The three nodes i, i', and i'' of s_f converge all three to p. We obtain a surface σ, bounded by s_i and s_f formed by faces that meet along edges; four of these edges meet at the point p. We call p a *vertex* of the surface σ. At the vertex p,

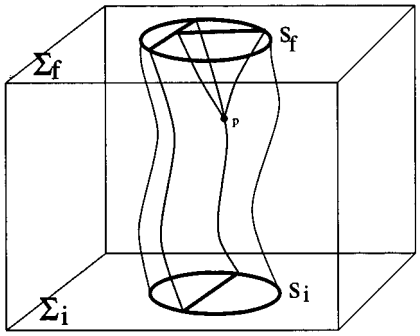


FIG. 3. Surface corresponding to a first order term.

σ branches. Notice that four edges and six faces meet in p, see Fig. 3.

We can imagine $\mathcal{M}$ as a spacetime and s_i as evolving continuously in a coordinate t from S_i to S_f. At the spacetime event p, the spin network branches: the node i generates the two new nodes i' and i'', which are born at i and then move away. A new face, spanned by the new edge that joins i' and i'', is born in p. The branching represents the elementary vertex of the theory, and is represented in Fig. 4.

The generalization of this construction to higher terms is immediate. A term of order n in T corresponds to a surface σ with n vertices. The (time) order in which the n D operators act determines an ordering for the vertices. An example of a term of order two is given in Fig. 5. It represents the transition from the s knot with two trivalent nodes connected by three links colored (3,5,7), to the s knot with the same graph, but colored (3,6,8). The intermediate step is the s-knot s_1, with four nodes.

In the construction we have described, each vertex has four adjacent edges. Some of these edges are generated by the nodes of the incoming s knot (the one at the right of the operator $\hat{D}$ corresponding to that vertex) and some by the outgoing s knot. At each node, we denote the first ones as ''past'' edges and the second ones as ''future'' edges. Thus, each edge emerges as a future edge from one vertex, or from the initial hypersurface, and ends as a past edge in another vertex, or in the final hypersurface. This defines a partial ordering of the vertices of each surface.

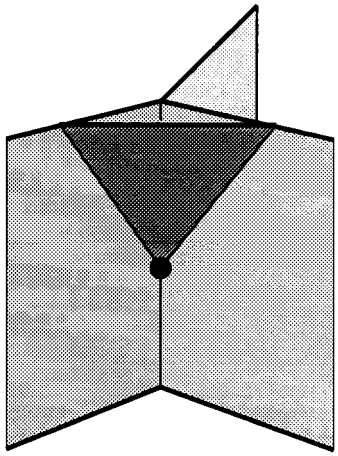


FIG. 4. The elementary vertex.

A short reflection will convince the reader that all the surfaces that we obtain satisfy the following property. Each face has the topology of a disk, and the ordered set vertices around a single face has at most one local maximum and most one local minimum. We say that a colored surface σ is ''well ordered'' if all its faces satisfy this property.

Now, observe the following. (i) The colored surface σ (with the vertices ordered) contains all the information needed to reconstruct the corresponding term in the expansion (55). In fact, the factors $A_\alpha(s)$ depend only on the coloring of the surface. (ii) Any well-ordered branched colored surface σ, with colorings satisfying Clebsch-Gordan conditions at the edges can be obtained from a term in Eq. (55).

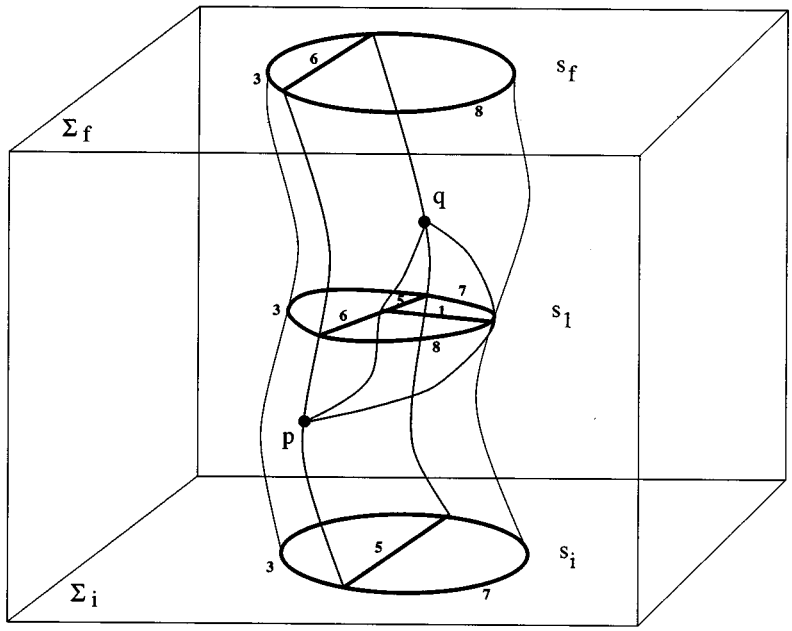


FIG. 5. A term of second order.

(iii) Two surfaces correspond to the same term if and only if there is a 4D diffeomorphism that sends one into the other.

These facts allow us to rewrite the expansion (55) as a sum over diffeomorphic inequivalent well-ordered surfaces σ bounded by s_i and s_f. Therefore we can write the propagator [see Eq. (35)] as a sum of terms labeled by topologically inequivalent branched well-ordered colored surfaces[9] s bounded by initial and final state:

$$P(s_f,s_i;T)=\sum_{\substack{\sigma \\ \partial\sigma=s_i\cup s_f}} \mathcal{A}[\sigma](T). \tag{56}$$

The weight $\mathcal{A}[\sigma](T)$ of the surface σ is given by a product over the $n(\sigma)$ vertices of σ:

$$\mathcal{A}[\sigma](T)=\frac{(-iT)^{n(\sigma)}}{n(\sigma)!}\prod_{v\in[\sigma]} A_v(\sigma). \tag{57}$$

The contribution $A_v(\sigma)$ of each vertex is given by the coefficients of the Hamiltonian constraint defined in Eq. (13):

$$A_v(\sigma)=A_\alpha(s). \tag{58}$$

(The nonvanishing matrix elements $\langle s'|D_\alpha|s\rangle$ have value 1.)

Can we attribute a physical meaning to the surfaces that enter the sum? The answer is yes. There is a natural way of interpreting a branched colored surface σ as a discrete (''quantum'') geometry. This geometrical interpretation was proposed in [8] in a slighty different context; it holds in the world sheet formulation of the simplicial model of GR [9]. First of all, consider a triangulation $\mathcal{T}$ of the manifold $\mathcal{M}$, and assume that the surface σ sits over the dual two-skeleton of the triangulation. As we shall see in Appendix B, this is the natural way of viewing the surfaces σ. Let a triangle (two cell) S of the triangulation $\mathcal{T}$ be punctured by the face f (say with color p) of σ in a point. Now, recall that according to canonical loop quantum gravity the colors of the spin networks are quanta of area: the area of a surface S pierced by a single link with color $p=2j$ is [18]

$$A(S)=16\pi\hbar G\sqrt{j(j+1)}. \tag{59}$$

In the spacetime picture, a link sweeps a 2D face f, which intersects S at a point. It is natural to suppose that the area of any spacetime two-surface S is similarly determined by the coloring of the world sheet. For instance, we may consider a three-dimensional hypersurface Σ that contains S, view the intersection between Σ and the colored surface σ as the ''instantaneous position of the spin-network state on the ADM time Σ,'' and assume that the results of the canonical theory can be applied. If we make this assumption, then we can say that the area of S is $A(S)$ given in Eq. (59). Therefore, a surface σ assigns a (possibly vanishing) area to each triangle of the triangulation $\mathcal{T}$. But fixing the areas of the triangles of a four-dimensional triangulation is equivalent to fixing a discretized four-geometry. Assigning areas is analogous to assigning the lengths of the links of the triangulation as in Regge calculus.[10] Thus, a surface σ defines a discretized four-geometry. The idea that areas of triangles could be variables more suitable than lengths of links in four dimensions was considered in [14,45]. Finally, more in general, we can say that a natural geometrical interpretation of the colors associated to the faces is the following: if a face has color $2j$, it contributes a quantum of area $16\pi\hbar G\sqrt{j(j+1)}$ to the area of each spacetime two-surface S at each point where it pierces S.

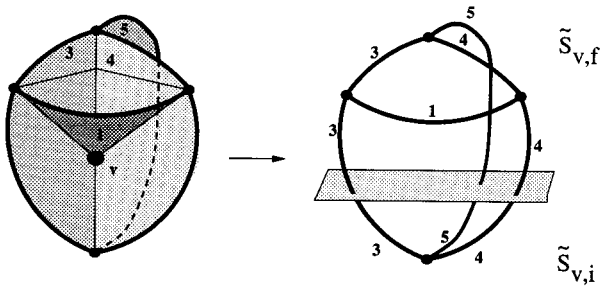


FIG. 6. The construction of the spin network S_v from the intersection of the surface with the boundary of a four-sphere surrounding the vertex. The spin network is then cut into its past and future components $\tilde{S}_{v,i}, \tilde{S}_{v,f}$.

This geometrical interpretation is ''natural,'' but not necessarily correct. In particular, the relation between the proper time T, and the spacetime geometry defined by the colors of the world sheet is not clear. This relation should be investigated before taking the geometrical interpretation too seriously.

Reconstruction of the $A_\alpha(S)$ coefficient from surface data. The coefficients $A_\alpha(s)$ can be reconstructed directly from the colored surface as follows. Let a vertex v have n_i past edges and n_f future edges. $A_v(\sigma)$ is nonvanishing only if $n_i=1$ and $n_f=3$ or if $n_i=3$ and $n_f=1$. In this case, $A_\alpha(s)$ is determined by the matrix elements of the Hamiltonian constraint.

It is very instructive to give an explicit construction of $A_v(\sigma)$. Consider a 4D neighborhood B of the vertex v. Consider the 3D boundary ∂B of B. Let S_v be the intersection between σ and ∂B. A short reflection will convince the reader that S_v is a colored graph in the 3D space ∂B, having $4=n_i+n_f$ nodes (that satisfy Clebsh-Gordan relations), corresponding to the intersections between the four edges emerging from p and ∂B, see Fig. 6.

Now, cut all the links of S_v that go from a past node to a future node. This procedure breaks S_v into two spin networks with (equal) open ends, which we denote as $\tilde{S}_{v,i}$ and $\tilde{S}_{v,f}$. (More precisely, these are s knots, because they are determined only up to diffeomorphisms.) The value of the vertex is given by the matrix elements of the Hamiltonian constraint between these two spin networks; namely,

$$A_v(\sigma)=\langle\tilde{S}_{v,f}|C[1]|\tilde{S}_{v,i}\rangle, \tag{60}$$

[9]For a general description of such surfaces and their properties, see, for instance, [44], and references therein.

[10]There is a difference: in order to define a geometry, the lengths of the links must satisfy certain inequalities. The areas of the triangles must satisfy certain inequalities, as well as some equalities among them; namely, they are not all independent.

see Eq. (11). This expression gives $A_v(\sigma)$ as a function of the colorings of the edges and faces adjacent to the vertex p. This function is universal, and characterizes general relativity, in the same manner in which the Feynman vertex factor characterizes a QFT. We compute the vertex function $A_v(S)$ explicitly for the simplest case in Appendix B. It turns out to be expressed in terms of SU(2) *n-j* symbols of the colorings.

Notice that it is the locality of the Hamiltonian constraint that allows the sum over surfaces construction. This is a peculiar form of locality, quite different from conventional QFT locality: The action is not local with respect to a background structure, but with respect to the spin networks themselves.

Equations (56)–(60) provide a definition of the proper time propagator of quantum general relativity as a topological sum over branched colored surfaces. They represent our main result.

V. CROSSING SYMMETRY

Above, we have considered a *reformulation* of loop quantum gravity as of a sum over surfaces. Here we propose a *modification* of the theory, suggested by the reformulation.

The value of the vertex $A_v(\sigma)$ that we have computed in the last section depends on two inputs. First, on the coloring of the edges and faces adjacent to the vertex v. Second, on the distinction between ''past'' and ''future'' edges, namely, on the way the vertex is located and oriented within the surface σ. We suspect that the appearence of this orientation dependence is a sign that something has gone wrong in the definition of the theory. The action of GR is local and 4D diff invariant, and therefore the action of the four-geometry of a small region (''the vertex'') is independent of how this region is sliced by equal time slices.[11]

Thus, we propose a modification of the theory in which the orientation dependence is removed. We say, in general, that in a theory defined by a sum over branched colored surfaces, with weights given by products of vertex factors, the vertex is ''crossing symmetry'' if its value depends on the adjacent colorings only, and not on the distinction between past and future edges. BF theory [11] and simplicial GR [8,9] are theories of this kind, and have crossing symmetric vertices. In this section we study the modification of the geometry of the vertex required to make it crossing symmetric.

We then say that a Hamiltonian constraint operator H has crossing symmetry if it defines a crossing-symmetric vertex via Eq. (60). The modification of the vertex that we consider in this section might be obtained from a different factor ordering of the Hamiltonian constraint and, as we shall see below, is strictly related 4D diff invariance. Thus, here we are exploring the idea that 4D diff invariance might fix residual factor ordering ambiguities. Of course, it should not be surprising that a spacetime formalism could simplify the discussion of 4D diff invariance, a notoriously tricky issue in the Hamiltonian framework.

Let us make clear that we present crossing symmetry only as a proposal to be explored. We do not have a rigorous derivation of crossing symmetry from first principles, but only a heuristic plausibility argument, which we better detail below.

Consider the path integral that formally defines $U(T)$ in a proper time gauge, namely in a gauge in which the lapse is spacially constant. Consider a four-metric g, that contributes to this path integral and a small spacetime region R, and let g_R be the restriction of g to R. The region R is sliced by the proper time slicing. Let $A_R(g)$ be the exponential of the action of this region. If the region is small enough, we can think of $A_R(g)$ as the matrix element of the evolution operator between ''before R'' and ''after R,'' where ''before'' and ''after'' are determined by the proper time slicing, and thus identify $A_R(g)$ with the vertex $A_v(\sigma)$. Now consider a different four-metric g' in the integral, containing a region R', such that $g'_{R'}$ is isometric to g_R, but sliced in a different manner by the proper time slicing (the reader will easily convince himself that such a metric exists in general). Since the action is local and 4D diff invariant, the contribution of g_R to the sum must be equal to the contribution of $g'_{R'}$, namely, $A_R(g)=A_{R'}(g')$. This implies that the matrix elements of the proper time Hamiltonian between ''before'' and ''after'' according to one slicing of R ought to be the same as the matrix elements of between ''before'' and ''after'' according any other slicing. In other words, the matrix elements should be invariant under a 4D rotation of R that changes what is before and what is after. If we require the same to hold in our sum over surfaces, we obtain the requirement that vertices be crossing symmetric.

This discussion shows that there is a relation between 4D diff invariance and crossing symmetry, because a 4D diffeomorphism ''rotates'' the vertex in 4D. Recall that the 4D diff invariance of the classical theory is expressed by the Poisson brackets

$$\{C[N],C[M]\}=C[N\vec{\partial}M-M\vec{\partial}N]. \tag{61}$$

One of the hard problems of the Hamiltonian quantization program is to define a quantization of the Hamiltonian constraint yielding a 4D diff-invariant quantum theory. In particular, implementation of 4D diff invariance is presumably the missing ingredient for fixing quantization ambiguities of the Hamiltonian constraint. Recall that the ambiguity in the definition of $C[N]$ was fixed in [24] and [25] to a large extent arbitrarily. Full implementation of the quantum version of Eq. (61) should ensure 4D diff invariance, but has proven hard to realize. We are therefore lead to the suggestion that we can cure the slicing dependence by taking advantage of the remaining operator ordering ambiguity, and at the same time, cure the excessive ordering ambiguity by imposing slicing independence. In other words, we can *impose some form of 4D diff-invariance requirement in order to reduce quantization ambiguity*. Here we are suggesting that in a covariant formalism crossing symmetry might be the key for implementing 4D diff invariance.

Vertices with crossing symmetry. Let us investigate the meaning and consequences of requiring crossing symmetry. First, we should require that different ways of cutting S_v with three nodes on one side and one node on the other yield the same $A_v(\sigma)$. This leads to conditions on the $A_\alpha(S)$ co-

[11]We are dealing with the Euclidean theory, so there is no light-cone structure that defines local notions of past and future.

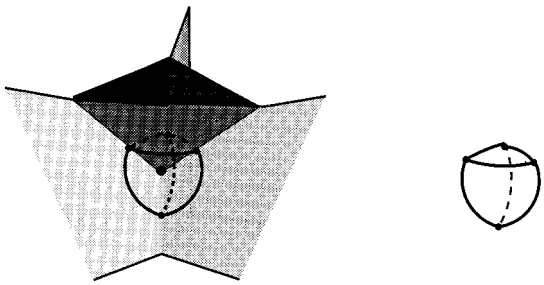


FIG. 7. The elementary vertex and its associated spin network S_v.

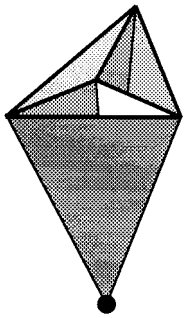


FIG. 9. The (0,4) transition: creation of a tetrahedron.

efficients, that will be studied elsewhere. A more interesting case is the following. First, let us help intuition by redrawing the elementary vertex of the theory (Fig. 4) in a more symmetric way. This is done in Fig. 7. (For simplicity, we restrict the following analysis to trivalent nodes.)

There are five topologically inequivalent ways of cutting S_v, giving, respectively, (n_i, n_f) (number of initial and final nodes) equal to (0,4), (1,3), (2,2), (3,1), and (4,0). The last two are the time reversal of the first two, leaving three genuinely independent cases. In Fig. 8, we show the possible cuts, and the corresponding spin networks transitions in the Hamiltonian picture. Time reversed cuts give just the opposite transitions.

Case (1,3) is the one described in the previous section (Fig. 6). Crossing symmetry requires that the Hamiltonian constraint generate the transitions (0,4) and (2,2)—described in the last column of Fig. 8—as well, *with the same amplitude*.

Consider these two new transitions. We begin with (0,4) (second line in Fig. 8). This transition represents a matrix element of a Hamiltonian that creates a ''small'' tetrahedron from the state with no loops. The fact that 4D invariance requires the presence of such ''birth'' terms has already been argued, on general grounds, in [46]. In terms of surfaces, the term looks as in Fig. 9.

This is *the very same surface as in Fig. 4, and in Fig. 7*, but drawn with a different orientation in ''spacetime.'' With this orientation, it describes a tetrahedral spin network emerging from nothing.

Can such a term originate from an ordering of the Hamiltonian constraint? Surprisingly, the answer is positive. We sketch here a hand waving argument. The regularized Hamiltonian constraint (FEE) is formed of two parts: a ''small loop'' that corresponds to the classical curvature (F) term and the term that ''grasps,'' the two hands of the T^2 operator in [17,24] (or the volume operator in [25]), corresponding to the triads EE (or the triads multimplied by a suitable density factor). Traditionally, the order chosen is FEE: the ''small loops'' are added *after* the grasping. Reverse this order, choosing EEF, and have the small loop being inserted first. Then the Hamiltonian constraint has nonvanishing action on the vacuum as well, because the grasping term can grasp the ''small loop.'' In particular, this may create a ''small'' tetrahedron. For instance, in the construction in [17,24], the T^2 can grasp itself, producing, precisely, a tetrahedron, see Fig. 10.

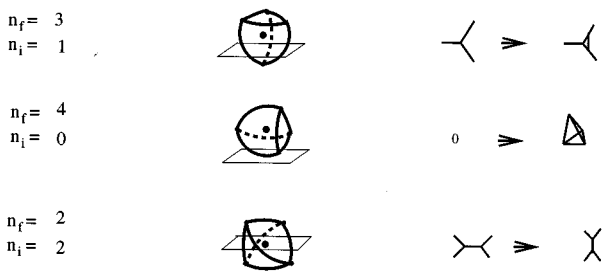


FIG. 8. The (1,3), (0,4), and (2,2) cuts of the elementary vertex, and, in the last column, the corresponding spin network transitions in the Hamiltonian picture. [For (3,1) and (4,0), look at (1,3) and (0,4) upside down, and reverse the arrow of the transition.]

The (2,2) term (third line of Fig. 8) gives a rearranging of two nodes. The corresponding surface looks as in Fig. 11. Again, this is just a different orientation of the same elementary vertex.

Now, we could search for an ordering of $C[N]$ yielding a vertex having crossing symmetry. But this task is superfluous since we already know what we should obtain. We can directly *postulate* that the Hamiltonian constraint yield crossing symmetry, and deduce the amplitudes of the (0,4) and (2,2) matrix elements from the value of the (1,3) vertex.[12]

There is a physical motivation that supports the above argument. It has been observed [47] that the orderings of the Hamiltonian constraint studied so far generate a dynamical evolution that appear to be excessively ''local.'' They preserve the general structure of the network on which they act, simply ''dressing'' nodes. These difficulties have been recently detailed in Ref. [6], where it is argued that no long range interaction is likely to emerge from a Hamiltonian with these features. As pointed out by Thiemann, the argument is far from conclusive, because it contains a jump from the nonphysical (gauge) coordinate evolution to the physical one, and this jump may be ungranted. In any case, adding the new vertices (2,2) and (0,4) would cure these potential difficulties. Also, we note that one of the consequences of adding the new vertices is that the faces of the surfaces in the sum do not need anymore to be (topologically) two disks, as fol-

[12]Notice that the symmetrization of the Hamiltonian constraint given in Eq. (8) can be seen as a first step in this covariantization of the operator: it is equivalent to the postulate that the (3,1) cut has the same value as the (1,3) cut.

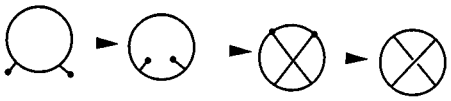


FIG. 10. Creation of a tetrahedron from the self-grasping of T^2.

lowed from the original Hamiltonian expansion, and there is no sense in the requirement of the surfaces being well-ordered.

Finally, notice that one could also search for the form of the vertices from general *a priori* requirements. An arbitrary crossing symmetric vertex is obtained by replacing Eq. (60) with a function $A(S_v)$ of the spin network S_v associated to the vertex

$$A_v(\sigma) \equiv A(S_v). \tag{62}$$

For instance, for trivalent nodes, S_v is a tetrahedron: the function $A()$ must respect tetrahedral invariance. Notice that there are not many functions with these features. A natural choice is

$$A(S_v) = Tet\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \tag{63}$$

where $a-f$ are the colors of the links of S_v and Tet is the totally symmetric form of the 6-j symbols ([48] and [21]). We think that a theory defined in this way is worth exploring.

VI. CONCLUSION

Our main result is contained in the equation

$$P(s_f, s_i; T) = \sum_{\substack{\sigma \\ \partial\sigma = s_i \cup s_f}} \frac{(-iT)^{n(\sigma)}}{n(\sigma)!} \prod_{v \in [\sigma]} A_v(\sigma). \tag{64}$$

[see Eqs. (56) and (57)], which expresses the dynamics of quantum general relativity in terms of a sum over surfaces σ.

More precisely, the proper time propagator of quantum GR can be expressed in terms of a sum over topologically inequivalent branched colored surfaces, bounded by the initial and final s knots. The contribution of each surface to the sum is the product of one factor per each vertex (branching point) of the surface. The contribution at each vertex is a simple SU(2)-invariant function $A_v(\sigma)$ of the colors of the faces and edges adjacent to the vertex. This function characterizes the quantum theory in the same manner in which the Feynman graph vertices characterize a quantum field theory. The vertex $A_v(\sigma)$ of general relativity is given by a product of Wigner 3n-j symbols [26].

The essential property of the expansion (64) is that it is finite order by order, and explicitly computable. This finiteness is intriguing. In order to calculate physical quantities, we must have the proper time propagator for multifingered proper times and we must integrate over the multifingered proper time. We expect that the integration could yield finite results if performed over expectation values of appropriate physical quantities. Work is in progress in this direction, and will be reported elsewhere. We close with the following comments.

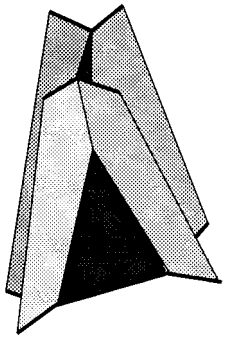


FIG. 11. The (2,2) vertex.

Our construction is strongly reminiscent of discretized quantum gravity on a lattice [20,49–51], particularly in its simplicial formulations [9,45]. It is shown in [9] that one can discretize general relativity over a simplicial lattice, and express the gravitational degrees of freedom as colored branched surfaces over the (dual) two-skeleton.[13] Even more remarkably, the partition function is given in the discretized case by a construction very similar to that given here: the contribution of a vertex is determined by the intersection between the boundary of a four-simplex around the branching point of the surfaces and the surface. This defines a spin network S_v, which, in the discretized case, can be any subgraph of the one-skeleton of a four-simplex. Therefore vertices have up to five edges and ten faces (see Appendix B) in the discretized case. In this paper, nonvanishing vertices have four edges and an arbitrary number of faces. Thus, the simplicial construction corresponds to a cut of the sum (56) in two respects: the maximum number of vertices is fixed by the triangulation, and vertices have ten faces at most.

One can view the sum over surfaces defined here as a version of Hawkings' integral over four-geometries. Indeed, a colored two surface defines a discrete four-geometry. The integral is replaced here by a sum, and explicit computation can be performed. Presumably, the construction can then be used to define a number of related theoretical tools such as partition functions, the Hartle-Hawking state, and others [2].

Each individual term in the expansion (55) is finite. Divergences can arise in summing the series, and in integrating over proper time.

The similarity with the formulation of string theory as a path integral over world sheets is tantalizing. On this, see the discussion in [7]. The *dynamics* is different. In string theory, the contribution of each surface to the sum is given by the area of the surface, and therefore it depends on a fixed background metric on the manifold. Here, on the contrary, the contribution of each surface depends only on the (coloring and) topology of the surface. Thus, quantum GR resembles a ''background-independent'' version of string theory. The techniques developed here could perhaps have relevance for connecting loop quantum gravity with string theory [7,53]—

[13]Surfaces seem to be playing an increasing role as a way to capture the gravitational field degrees of freedom. See for instance [52].

or for the construction of a nonperturbative background-independent formulation of string theory.

ACKNOWLEDGMENTS

We are particularly indebted to Jim Hartle, Gary Horowitz, and John Baez for extensive discussions during which various aspects of this work were clarified. We thank Don Marolf for key suggestions in an early stage of the work, Bernie Brügmann for criticisms and suggestions, Thomas Thiemann for a detailed and valuable critical reading of the manuscript, Roberto DePietri and Roumen Borissov, for pointing out and correcting several errors, and Lee Smolin, Seth Major, Jerzy Lewandowski, Louis Crane, John Baker, and Abhay Ashtekar for suggestions and discussions. This work was partially supported by NSF Grant Nos. PHY-5-3840400, PHY-9515506, PHY95-14240, and the Eberly research fund of PSU. C.R. thanks the A. Einstein Institute in Potsdam, and the Physics Department of the University of Roma for hospitality during the preparation of this work.

APPENDIX A: TERMINOLOGY

To help the reader, we collect here a list of terms employed.

Node: Point in 3D space where the links of a spin network meet.

Link: Line in 3D space connecting two nodes of a spin network.

Face: Surface in 4D spacetime (swept by a link).

Edge: Line in 4D spacetime where several faces meet (swept by a node).

Vertex: Point in 4D spacetime where several edges meet.

A spin network is formed by nodes and links. A branched surface is formed by faces, edges, and vertices. For the branched surfaces that live on the two-skeleton of the dual triangulation of the manifold in simplicial BF theory, faces, edges, and vertices live on two-, one-, and zero-cells, respectively, of the cellular decomposition dual to the simplicial triangulation. They are therefore associated to four-, three-, and two-simplices of the triangulation, respectively. Therefore a vertex corresponds to a four-simplex, an edge to a tetrahedron, and a face to (its dual) triangle.

APPENDIX B: COMPARISON WITH THE OOGURI-CRANE-YETTER 4D TQFT

The structure of quantum general relativity in the form presented in this paper is surprisingly similar to the Ooguri-Crane-Yetter (OCY) four-dimensional topological quantum field theory [10,11], a rigorously defined simplicial lattice version of four-dimensional SU(2) BF theory (see also [58]). More specifically, our expression for the proper time propagator $U(T)$ of GR as a sum over world sheets resembles in many ways the world sheet sum [8,48][14] for the projector on physical states U_{OCY} of the OCY model.

The OCY model is a 4D generalization of the Ponzano-Regge-Turaev-Viro (PRTV) model [12,13], which, in turn, can be seen as a quantization of 3D GR, or a quantization of 3D Chern-Simon theory. In [14] it was shown that the PRTV model is a theory of the dynamics of spin networks (loops in the terminology of [14]) having the same physical interpretation as the spin network basis states in continuum 3D GR. Thus one might expect a similarity between the kinematics of the OCY model and loop quantized GR [14]. On the other hand, the 4D OCY model as with the 3D PRTV model, but unlike 4D GR has no local degrees of freedom, so one also expects large differences between the theories.

In this section, we sketch the OCY theory, outline a construction of the world sheet sum for the partition function $Z_B F$ of the OCY model along the lines of [8],[15] and discuss its similarities and differences with the formulation of quantum GR presented here. We believe that this comparison helps illuminate the much debated issue of the relation between quantum gravity and TQFT's [15].

We introduce here Ooguri's original version [10] the OCY model heuristically, as a discretization of BF theory without cosmological constant. BF theory is given in terms of two fields, an SU(2) connection A^i, with curvature F^i, and an SU(2) algebra valued two-form B^i, by the action [54]

$$S_{\mathrm{BF}}=\int B^i\wedge F^i. \tag{B1}$$

Before proceeding, it is interesting to note that conventional general relativity can be obtained from BF theory by simply adding a constraint term. Indeed one can show that the theory

$$S_{\mathrm{GR}}=\int B^i\wedge F^i+\phi_{ij}B^i\wedge B^j, \tag{B2}$$

where the Lagrange multiplier ϕ_{ij} is traceless and symmetric, is equivalent to general relativity [55].

Consider the partition function of the BF theory

$$Z_{\mathrm{BF}}=\int [dA][dB]e^{-i\int B^i\wedge F^i}. \tag{B3}$$

Integrating over B, we obtain

$$Z_{\mathrm{BF}}=\int [dA]\delta[F]; \tag{B4}$$

namely, an integral over flat SU(2) connections. Let us define a lattice version of this theory by fixing a simplicial decomposition of the 4D manifold. (See also [51].)

Consider the dual of the simplicial decomposition. There is one element of this dual cellular decomposition that plays a central role in the construction: the ''wedge.'' Consider a dual-two-face f. A dual-two-face is a 2D poligon. It intersects a two-face of the simplicial decomposition in a ''cen-

[14]The world sheet sum of [48] is actually for the 3D Ponzano-Regge-Turaev-Viro (PRTV) model, but is easily extended to the OCY model. Iwasaki [50] has proposed an interesting alternative, a closely related world sheet formulation of the PRTV model which is also easily extended to the OCY model.

[15]The construction outlined here is a sort of baby version of that for simplicial GR in [46].

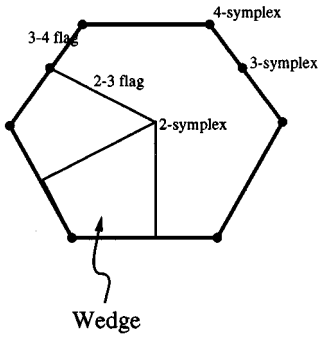


FIG. 12. A dual-two face and its decomposition in wedges.

tral'' point o. Its vertices are centers of four-symplices and its sides are lines connecting such centers. Each of these sides, which connects the centers of two-simplices, crosses the tetrahedron that forms the boundary between the two simplices. Let p be the crossing point. Each point p cuts one of the sides of the polygon f. By drawing lines connecting the points p to O, we divide the polygon f in quadrangles, called ''wedges.'' A wedge is thus a 2D quadrangle that has four sides: two of these are 1D lines that join centers of four-simplices with (the center p of) a bounding tetrahedron; these are denoted 3-4 flags. The other two lines join the center of a tetrahedron with the center of a two-face. These are denoted 3-2 flags, see Fig. 12.

We choose to represent the connection by means of group elements associated to 1D elements in the dual cellular decomposition. More precisely, we associate a group element U to each 2-3 flag (segment connecting the center of a two-simplex with the center of one of the tetrahedra surrounding it), and one group element W to each 3-4 flag (segment connecting the center of a tetrahedron with the center of an adjacent four-simplex). These group elements can be thought of as the exponential of the connection along the segments. Each wedge w is bounded by four such segments (two of the 2-3 kind and two of the 3-4 kind); let $U_1(w), U_2(w), W_1(w), W_2(w)$ be the group elements associated to the segments that bound the wedge w. We can express the requirement that the connection is flat by requiring that the holonomy of the connection around each wedge is trivial. Then a discretization of Eq. (B4) is given by

$$Z_{BF}=\int [dU][dW]\prod_w \delta[U_1(w)U_2(w)W_1(w)W_2(w)] \tag{B5}$$

where the δ function is the δ function of the unit on the SU(2) group. We can expand the δ function in characters. For each w, we have

$$\delta[U_1(w)U_2(w)W_1(w)W_2(w)] = \sum_j (2j+1)\mathrm{Tr}_j[U_1(w)U_2(w)W_1(w)W_2(w)], \tag{B6}$$

where j labels the irreducible representations of SU(2), and $\mathrm{Tr}_j(U)$ is the trace of the group element U in the representation j. Using this, we can rewrite Eq. (B5) as an integral over group elements U and W associated to segments and half integers j associated to wedges:

$$Z_{\mathrm{BF}}=\int [dU][dW]\sum_{[j]}\prod_w [2j(w)+1] \times \mathrm{Tr}_{j(w)}[U_1(w)U_2(w)W_1(w)W_2(w)]. \tag{B7}$$

We can view the group elements U and W as the discrete version of the connection A, the j's as a discrete version of the two form B, and the expression $\mathrm{Tr}_{j(w)}[U_1(w)U_2(w)W_1(w)W_2(w)]$ as the discrete version of the expression $\exp(-iB^i\wedge F^i)$.

Next, let us perform the group integration in Eq. (B7) explicitly. By integrating over the group elements U we force the colors of the wedges belonging to the same dual two-cell to be equal. By integrating over the group elements W we force the j's of the (four) dual two-cells that join on a one-cell to satisfy the Clebsch-Gordan relation where they meet, leaving an extra degree of freedom J associated to each such dual one-cell; J runs over the independent couplings of four SU(2) representations. Finally, we end up with numerical factors associated to the zero-cells of the dual triangulation (plus other factors associated to faces and edges, which we disregard here in order not to make the exposition too heavy). Such numerical factors turn out to be 15-j symbols associated to the five J's of the five one-cells and the ten j's of the ten two-cells adjacent to each vertex.[16] Performing these integrations explicitly is a simple and interesting exercise. After these integrations over the group elements, the theory is therefore reduced to a sum over colorings on the two-cells and one-cells, satisfying Clebsch-Gordan relations. We can interpret a zero color as no surface at all, and identify the two-cells with faces and the one-cells with edges of branched colored surfaces. Thus, we can write the partition function as a sum over branched colored surfaces living on the dual two skeleton of the simplicial triangulation. We obtain

$$Z_{\mathrm{BF}}=\sum_\sigma A_{\mathrm{BF}}[\sigma] \tag{B8}$$

where the contribution of each surface is (up to the face and edge factors we have disregarded for simplicity) a product of vertices's factors

$$A_{\mathrm{BF}}[\sigma]=\prod_v A_{\mathrm{BF},v}(\sigma); \tag{B9}$$

the vertex factor is the 15-j symbol of the colorings adjacent to the vertex.

The similarity of this result with the construction in this paper, Eqs. (56), (57), (58), is striking. In both cases, we have a sum over the same kind of branched colored surfaces, and the weight for each surface is the product of vertex fac-

[16] A dual zero-cell is always adjacent to five one-cells and ten two-cells, because it corresponds to a four-simplex of the original simplicial decomposition, which is bounded by five tetrahedra and ten faces.

tors, where vertex factors are simple SU(2)-invariant functions of the adjacent colorings. Thus, the structure of quantum general relativity turns out to be extremely similar to the structure of a topological quantum field theory. Of course there are differences, and these differences are crucial. Let us examine them in detail.

First of all, the world sheet amplitudes in the BF theory that we are considering are the amplitudes in the projector on physical states, while the GR world sheet amplitudes of the present paper are from the sum for the proper time propagator $U(T)$, so we might be comparing apples and oranges. However, if we accept the not unreasonable hypothesis that $U(T)$ is a partial sum of terms in a sum over surfaces for U in GR, we can compare the theories in a direct way.

The vertex factor is different in the two theories: in BF theory it is a 15-j symbol, while in GR it is a combination of 9-j and 6-j symbols. This difference depends on the different dynamics of the two theories, and should be at the root of the other differences.

In the case of BF theory there is a crucial theorem holding: triangulation independence. Refining the triangulation does not change the overall sum. This is the reason for which the theory is topological, and is a consequence of the fact that the classical theory has no local degrees of freedom. In GR, nothing similar holds, because GR has genuine local (*although nonlocalized*) degrees of freedom. Therefore there is no reason to expect anything similar to triangulation independence for GR.

The ensemble of surfaces over which the sum is defined is different in the two cases. In the BF case, we sum over surfaces over a fixed triangulation. In the case of GR, we sum over all topologically inequivalent surfaces, with an arbitrary number of vertices. Therefore, in the case of BF theory the surfaces to be considered are finite in number. In the case of GR we have to sum over arbitrarily complicated surfaces or, equivalently, sum over arbitrarily fine triangulations of the manifold as well. Notice that this difference is a consequence of the previous point, namely, triangulation independence of BF. We could average over arbitrarily fine triangulations in BF as well, but this would not affect the result, because each triangulation yields the same contribution as the coarsest one. Therefore, *diff invariance of the sum is implemented in two different ways in BF and in GR, corresponding to the fact that BF is topological, while GR is not: in BF, invariance is obtained thanks to triangulation independence; in GR invariance is obtained by summing over arbitrarily fine triangulations.*

The rigorous version of BF theory requires SU(2) to be replaced by quantum SU(2). This can be seen simply as a smart stratagem for regularizing the sum in an invariant manner, yielding a finite result. Notice that in GR regularizing SU(2) to quantum SU(2) would not guarantee an overall finite sum, because the surfaces themselves are infinite in number. Thus, quantum GR does not admit a rigorous finite version as quantum BF, at least as far as we can presently see, even if one attempts to replace SU(2) with quantum SU(2) in GR [56].

The vertices of BF have always five edges and ten faces, while vertices of GR have (at least with the ordering considered so far) four edges and an arbitrary number of faces.

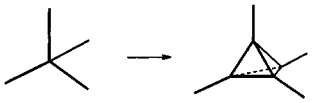


FIG. 13. The elementary vertex of BF theory. Notice that there are 5 nodes and 10 links, yielding 15 colors. The value of this vertex is the 15-j symbol of their 15 colors.

These points illuminate the difference between quantum GR and topological field theories. Let us discuss this point in more detail.

Both theories are invariant under diffeomorphism. However, diffeomorphism invariance does not imply that a quantum theory is topological in the sense of having a finite number of degrees of freedom. We expect GR to have an infinite number of degrees of freedom. Thus Atiyah's axioms for topological quantum field theory are likely to be suitable for quantum general relativity as well, if we drop the request that the Hilbert spaces attached to boundaries of the four-manifold be finite dimensional.

Finally, we may turn the comparison the other way around, and describe quantum BF theory in the language so far used for quantum GR. We can capture quantum BF theory in terms of its vertex. A BF vertex has five edges and ten faces. Assume that one of these edges comes from the past, and four go to the future (the other cases are given by crossing symmetry, that clearly holds in BF theory). A moment of reflection shows that the elementary vertex of BF theory ''opens up'' a four-valent intersection of a spin-network into a ''small'' tetrahedron, see Fig. 13.

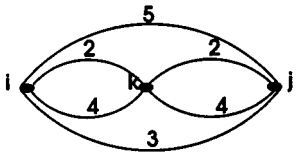
The matrix element of the Hamiltonian between these two (partial) spin networks is the 15-j symbol of the 15 colors associated to the four links and the one node of the incoming spin network, and the six other links and four other nodes of the newly created tetrahedron. It would be interesting to derive this Hamiltonian from a Hamiltonian loop quantization of BF theory.

APPENDIX C: DIFF INVARIANT SCALAR PRODUCT

We work out here an example of diff-invariant scalar product between s-knot states [32]. (On the inner product between spin networks see [29,57], and [21] which we follow here.) Let s be the s knot defined as follows. s has three four-valent nodes, i, j, and k, and the six links

$$(ki,2),(kj,2),(ki,4),(kj,4),(ij,3),(ij,5), \tag{C1}$$

where each link is indicated by the two nodes it connects and its color. Explicitly,

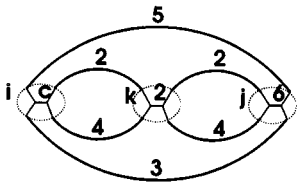


(C2)

To specify the state, we have to give also the coloring of the nodes. We choose an expansion of the nodes in a trivalent

graph (in each node) by pairing the two links colored 2 and 4. We have one virtual link for every node (we denote them e_i, e_j and e_k), which we assume to be colored as $e_i{:}c, e_j{:}6$ and $e_k{:}2$, where c (which can take the values 6, 4, and 2) will be specified later on:

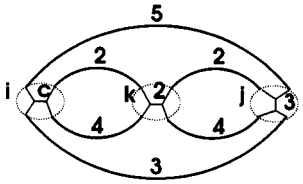
s =



(C3)

Next, let us define the s knot s'. Let it be the same as above, but with a different coloring of the node j. We expand j by pairing $(kj,2)$ with $(ij,5)$ and $(kj,4)$ with $(ij,3)$. Let the internal link have color 3:

s' =



(C4)

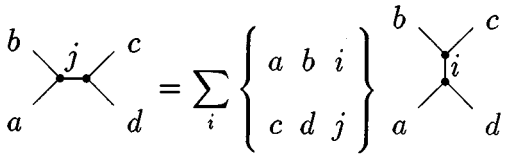
Let us compute the scalar product $\langle s|s'\rangle$. First, we have to list the automorphisms of the spin network (taking link colors, but not node colors into account). There is only one nontrivial automorphism α: it exchanges i and j. Thus, Eq. (5) gives

$$\langle s|s'\rangle = \langle S|S'\rangle + \langle \alpha S|S'\rangle, \tag{C5}$$

where $S \in s$ and $S' \in s'$ have been selected to have the same graph, with the same colored links. The only contribution to Eq. (C5) comes from the nodes. The node k is the same in the two states and therefore gives no contribution (recall we have chosen normalized states). Thus, we have

$$\langle s|s'\rangle = \langle S|S'\rangle_i \langle S|S'\rangle_j + \langle \alpha S|S'\rangle_i \langle \alpha S|S'\rangle_j, \tag{C6}$$

where we have indicated with $\langle|\rangle_i$ the scalar product restricted to the space of one node. Spin-network states with the same trivalent expansion are orthonormal. The change of basis is given [21] by the recoupling theorem



$$= \sum_i \begin{Bmatrix} a & b & i \\ c & d & j \end{Bmatrix}$$

(C7)

where the quantities $\{^{abi}_{cdj}\}$ are su(2) six-j symbols (normalized as in [48]). This gives us immediately

$$\langle S|S'\rangle_i = 1, \tag{C8}$$

$$\langle S|S'\rangle_j = \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & 6 \end{Bmatrix}, \tag{C9}$$

$$\langle \alpha S|S'\rangle_i = \delta_{c6}, \tag{C10}$$

$$\langle \alpha S|S'\rangle_j = \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & c \end{Bmatrix}. \tag{C11}$$

Therefore

$$\langle s|s'\rangle = \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & 6 \end{Bmatrix} + \delta_{c6} \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & c \end{Bmatrix}. \tag{C12}$$

Thus, if $c = 6$ there are two contributions to the scalar product, one from each of the two elements of the automorphism group of the spin network and we have

$$\langle s|s'\rangle = 2 \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & 6 \end{Bmatrix} = \frac{112}{75}. \tag{C13}$$

While if $c = 2$ or $c = 4$,

$$\langle s|s'\rangle = \begin{Bmatrix} 4 & 2 & 3 \\ 5 & 3 & 6 \end{Bmatrix} = \frac{56}{75}. \tag{C14}$$

[1] For an overview, see C. Isham, gr-qc/9510063.
[2] J. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983); S. W. Hawking, in *Relativity, Groups and Topology, Les Houches Session XL*, edited by B. DeWitt and R. Stora (North Holland, Amsterdam, 1984).
[3] R. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
[4] C. Teitelboim, Phys. Lett. **96B**, 77 (1980).
[5] R. Feynman, Phys. Rev. **80**, 440 (1950); Y. Nambu, Prog. Theor. Phys. **5**, 82 (1950).
[6] L. Smolin, gr-qc/9609034.
[7] J. Baez, in *Knots and Quantum Gravity*, edited by J. Baez (Oxford University Press, Oxford, 1994).
[8] M. Reisenberger, gr-qc/9412035.
[9] M. Reisenberger, gr-qc/9609002, (in preparation).
[10] H. Ooguri, Mod. Phys. Lett. A **7**, 2799 (1992).
[11] L. Crane and D. N. Yetter, in *Quantum Topology*, edited by R. A. Baadhio and L. H. Kauffman (World Scientific, Singapore, 1993); L. Crane, L. H. Kauffman, and D. N. Yetter, in *ibid.*, pp. 131–138. See also [58].
[12] G. Ponzano and T. Regge, in *Spectroscopy and Group Theoretical Methods in Physics*, edited by F. Bloch (North Holland, Amsterdam, 1968); H. Ooguri and N. Sasakura, Mod. Phys. Lett. A **6**, 3591 (1991); H. Ooguri, Nucl. Phys. **B382**, 276 (1992).
[13] V. G. Turaev and O. V. Viro, Topology **21**, 865 (1992).
[14] C. Rovelli, Phys. Rev. D **48**, 2702 (1993).
[15] L. Crane, J. Math. Phys. (N.Y.) **36**, 6180 (1995); J. Baez, *ibid.* **36**, 6073 (1995); M. Atiyah, *ibid.* **36**, 6069 (1995); A. S. Cattaneo, P. Cotta-Ramusino, J. Frölich, and M. Martellini, *ibid.* **36**, 6137 (1995); J. Barret, *ibid.* **36**, 6161 (1995); L. Smolin,

ibid. **36**, 6417 (1995); L. Crane, in *Knots and Quantum Gravity* [7].

[16] J. B. Hartle, in *Gravitation and Quantization, Les Houches, Session LVII,* edited by B. Julia and J. Zinn-Justin (Elsevier Science, New York, 1995).

[17] C. Rovelli and L. Smolin, Phys. Rev. Lett. **61**, 1155 (1988); Nucl. Phys. **B331**, 80 (1990). For a recent introductive review, see C. Rovelli, Helv. Phys. Acta **69**, 582 (1996), and complete references therein. For a detailed introduction to the formalism needed to perform the calculations in the scheme developed here, see [21]. General introductions to the loop formalism can be found in C. Rovelli, Class. Quantum Grav. **8**, 1613 (1991); R. Gambini and J. Pullin, *Loops, Knots and Quantum Gravity* (Cambridge University Press, Cambridge, England, 1997); A. Ashtekar, in *Gravitation and Quantization, Les Houches, Session LVII* [16].

[18] C. Rovelli and L. Smolin, Nucl. Phys. **B442**, 593 (1995).

[19] A. Ashtekar and J. Lewandowski, J. Geom. Phys. **17**, 191 (1995); J. Lewandowski, talk given at the *Workshop on Canonical and Quantum Gravity*, Warsaw, 1995 (unpublished).

[20] R. Loll, Phys. Rev. Lett. **75**, 3048 (1995); Nucl. Phys. **B444**, 619 (1995); **B460**, 143 (1996); Phys. Rev. D **54**, 5381 (1996).

[21] R. DePietri and C. Rovelli, Phys. Rev. D **54**, 2664 (1996).

[22] S. Frittelli, L. Lehner, and C. Rovelli, Class. Quantum Grav. **13**, 2921 (1996).

[23] J. Lewandowski, Class. Quantum Grav. **14**, 71 (1997).

[24] T. Jacobson and L. Smolin, Nucl. Phys. **B229**, 295 (1988); V. Husain, *ibid.* **B313**, 711 (1988); W. P. Blencowe, *ibid.* **B341**, 213 (1990); B. Brügman and J. Pullin, *ibid.* **B341**, 221 (1990); B. Brügman, R. Gambini, and J. Pullin, Phys. Rev. Lett. **68**, 431 (1992); Nucl. Phys. **B385**, 587 (1992); Gen. Relativ. Gravit. **25**, 1 (1993); C. Rovelli and L. Smolin, Phys. Rev. Lett. **72**, 446 (1994); R. Gambini and J. Pullin, Phys. Rev. D **54**, 5935 (1996); Class. Quantum Grav. **13**, L125,446 (1996); gr-qc/9703042; R. Borissov, Phys. Rev. D **55**, 2059 (1997); **55**, 6099 (1997); K. Ezawa, gr-qc/9601050.

[25] T. Thiemann, Phys. Lett. B **380**, 257 (1996); gr-qc/9606089; gr-qc/9606090.

[26] R. Borrisov, R. DePietri, and C. Rovelli, Class. Quantum Grav. (to be published).

[27] J. Lewandowski, Int. J. Mod. Phys. D **3**, 207 (1994); C. Rovelli and L. Smolin, Phys. Rev. D **52**, 5743 (1995). L. Smolin, gr-qc/9702030.

[28] J. Baez, Adv. Math. **117**, 253 (1996); in *The Interface of Knots and Physics,* edited by L. H. Kauffman (American Mathematical Society, Providence, RI, 1966).

[29] T. Thiemann, hep-th/9601105.

[30] A. Ashtekar and J. Lewandowski, in *Knots and Quantum Gravity* (Clarendon, Oxford, 1994); A. Ashtekar and C. Isham, Class. Quantum Grav. **9**, 1433 (1992); A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao, and T. Thiemann, J. Math. Phys. (N.Y.) **36**, 6456 (1995).

[31] R. DePietri, Class. Quantum Grav. **14**, 53 (1997).

[32] D. Maroff, Ph.D. dissertation, 1992; Class. Quantum Grav. **12**, 1199 (1995); D. Marolf, *The Spectral Analysis Inner Product*, Proceedings of the VIIth Marcel Grossman Conference, edited by R. Ruffini and M. Keiser (World Scientific, Singapore, 1995).

[33] A. Higuchi, Class. Quantum Grav. **8**, 1983 (1991).

[34] J. F. Barbero, Phys. Rev. D **51**, 5507 (1995); T. Thiemann, Class. Quantum Grav. **13**, 1383 (1996); A. Ashtekar, Phys. Rev. D **53**, 2865 (1996).

[35] C. Rovelli, J. Math. Phys. (N.Y.) **36**, 5629 (1995).

[36] B. Brügmann, Nucl. Phys. **B474**, 249 (1996).

[37] J. Lewandowski (private communication).

[38] C. Rovelli, Class. Quantum Grav. **8**, 297 (1991).

[39] J. D. Brown and K. Kuchar, Phys. Rev. D **51**, 5600 (1995).

[40] J. Lewandowski (private communication).

[41] N. Grot and C. Rovelli, J. Math. Phys. (N.Y.) **37**, 3014 (1996).

[42] J. A. Zapata, gr-qc/9703037; gr-qc/9703038.

[43] E. P. T. Liang, Phys. Rev. D **5**, 2458 (1972); C. J. Isham, Proc. R. Soc. London **A351**, 209 (1976); C. Teitelboim, in *An Einstein Centenary Volume*, edited by A. Held (Plenum, New York, 1980).

[44] J.S. Carter and M. Saito, in *Knots and Quantum Gravity* [7].

[45] R. M. Williams, J. Math. Phys. (N.Y.) **36**, 6276 (1995).

[46] M. Reisenberger, Nucl. Phys. **B457**, 643 (1995).

[47] B. Brügmann (private communication).

[48] L. H. Kauffman and S. L. Lins, *Temperley-Lieb Recoupling Theory and Invariant of 3-Manifolds* (Princeton University Press, Princeton, 1994).

[49] H. Fort, R. Gambini, and J. Pullin, Phys. Rev. D **56**, 2127 (1997); G. Immirzi, Class. Quantum Grav. **13**, 2385 (1996).

[50] J. Iwasaki, J. Math. Phys. (N.Y.) **36**, 6288 (1995).

[51] J. A. Zapata, Class. Quantum Grav. **13**, 2617 (1986).

[52] S. Frittelli, E. T. Newman, and C. N. Kozameh, J. Math. Phys. (N.Y.) **36**, 6397 (1995).

[53] L. Smolin, gr-qc/9609031.

[54] A. S. Schwarz, Lett. Math. Phys. **2**, 247 (1978); Commun. Math. Phys. **67**, 1 (1979); G. Horowitz, *ibid.* **125**, 417 (1989); M. Blau and G. Thompson, Ann. Phys. (N.Y.) **209**, 129 (1991).

[55] J. F. Plebanski, J. Math. Phys. (N.Y.) **18**, 2511 (1977); T. Jacobson and L. Smolin, Phys. Lett. B **196**, 39 (1987); Class. Quantum Grav. **5**, 583 (1988); R. Capovilla, J. Dell, and T. Jacobson, *ibid.* **8**, 41 (1991); **8**, 59 (1991); R. Capovilla and T. Jacobson, Mod. Phys. Lett. A **7**, 1871 (1992).

[56] R. Borissov, S. Major, and L. Smolin, Class. Quantum Grav. **13**, 3183 (1996).

[57] J. C. Baez, Adv. Math. **117**, 253 (1996); gr-qc/9411007; in *The Interface of Knots and Physics* [28]; gr-qc/9504036.

[58] L. Crane and I. Frenkel, J. Math. Phys. (N.Y.) **35**, 5136 (1994); L. Crane and D. N. Yetter (unpublished); J. Knot Theory Ramif. (to be published); L. Crane, L. H. Kauffman, and D. N. Yetter, *ibid.* **6**, 177 (1997); P. Cotta-Ramusino and M. Martellini, in *Knots and Quantum Gravity* [7].