Loop quantum gravity and discrete space-time

Gordon McCabe

August 21, 2018

Abstract

Loop quantum gravity purports to be a viable candidate for a physical theory which quantizes general relativity, and captures the discrete nature of space-time at a fundamental level. This paper subjects that joint claim to critical scrutiny. The paper begins by reviewing the canonical formulation, proceeds to an exposition of Ashtekar's 'new variables', and then details the development of loop quantum gravity, focusing on the definition of spin-network states and area operators. The paper concludes with a critical analysis of loop quantum gravity's credentials as a physical theory. A number of the widely-held beliefs, particularly those related to the discrete nature of space, are rejected as misconceptions.

1 Introduction

Loop quantum gravity (LQG) is considered by many to be the best prospect for a mathematically rigorous, empirically adequate, and conceptually well-defined theory of quantum space-time geometry. The architects and engineers of the theory proudly claim to have identified fundamental discrete geometrical structures, together with area operators and volume operators that possess discrete spectra. The smooth and continuous manifold structure of general relativity is commonly relegated to a mere semi-classical limit, and the proponents of LQG talk of having discovered the 'atoms' or 'grains' of space.

But are these claims justified? Is the theory well-conceived, or has the LQG community injected unwarranted philosophical preconceptions into its foundations? Has the desire for mathematical tractability derailed the theory, divesting it of physical content?

In the first section we begin by reviewing the canonical formulation of general relativity, and the original 'metric-based' attempt to quantize it. In the process, we will discuss the distinction between local Lorentz covariance and general covariance, a distinction which will become highly pertinent in our later assessment of LQG. We will also briefly discuss the problem of time.

In the second section we will introduce Ashtekar's new variables, and carefully scrutinise Barbero's modification of it. The third section provides an exposition of Loop Quantum Gravity, explaining how the kinematic state space is obtained, along with the self-adjoint operators which purportedly represent area and volume. In the process, the role of spin-network states is explained.

The paper concludes with a critical analysis of the development of Loop Quantum Gravity, and its credentials as a physical theory. A number of the cherished beliefs held by the LQG community, particularly those related to the discrete nature of space, are shown to be false or unjustified.

2 Canonical Quantum Gravity

Let's begin by briefly reviewing the original 'metric-based' approach to the canonical quantization of gravity.

In the lexicon of physics, a 'canonical' formulation is one which uses the techniques of Hamiltonian mechanics. In the canonical approach to classical general relativity, one does not deal directly with a 4-dimensional space-time, but with the evolution of fields upon a fixed 3-manifold Σ. In the metric-based formulation, these fields are the metric tensor γ_{ij}, the conjugate momentum π^{kl}, and any matter fields ϕ. One deals with a configuration space, a phase space, a Hamiltonian, constraint equations, and dynamical evolution equations.

Whilst a Riemannian metric tensor γ_{ij} upon a 3-manifold describes the geometrical configuration of space, the rate of change of the geometry is specified by the extrinsic curvature tensor, K_{ij}, otherwise known as the second fundamental form. Whilst K_{ij} is the analogue of velocity, the analogue of momentum is the tensor field π^{kl}, defined as

$$\pi^{kl} = \gamma^{1/2}(\gamma^{kl}K^i_i - K^{kl})$$

where $\gamma^{1/2} \equiv (\det\gamma)^{1/2}$.

The set of all possible pairs (γ_{ij}, π^{kl}) is essentially the phase space Γ of canonical general relativity. However, not every point of a general relativistic phase space constitutes permissible initial data. A pair (γ_{ij}, π^{kl}) must satisfy a scalar constraint and a vector constraint to constitute acceptable initial data. Denote the scalar constraint function as $H_0(x; \gamma, \pi]$, and the vector constraint function as $H_i(x; \gamma, \pi]$.[1]

The constraint equations are related to the dynamics of general relativity. The dynamics require one to specify a (possibly time-dependent) lapse scalar field $N(t)$, and a (possibly time-dependent) shift vector field $N^i(t)$ on the 3-manifold Σ. The lapse and shift can be freely specified, but once a choice has been made, one has a (possibly time-dependent) function on the phase space

[1]Semicolons are used in these expressions to separate the point arguments x from those arguments which are fields themselves, i.e. γ and π. Curved brackets are used to enclose the point arguments, and angular brackets for the fields.

which plays the role of the Hamiltonian $H_{(N,N^i)}$:

$$
\begin{aligned}
H_{(N,N^i)}(t)[\gamma,\pi] &= \int_{\Sigma_t} N(x,t)H_0(x;\gamma,\pi] + N^i(x,t)H_i(x;\gamma,\pi]\Omega_\gamma \\
&= \int_{\Sigma_t} N(x,t)H_0(x;\gamma,\pi]\Omega_\gamma + \int_{\Sigma_t} N^i(x,t)H_i(x;\gamma,\pi]\Omega_\gamma \\
&= H_N[\gamma,\pi] + H_{N^i}[\gamma,\pi]
\end{aligned}
$$

$H_{(N,N^i)}$ is defined upon the entire phase space Γ, but vanishes upon the constraint submanifold. This Hamiltonian $H_{(N,N^i)}$, in conjunction with a symplectic structure $\Omega_{\alpha\beta}$ on phase space, and the corresponding Poisson bracket, yields the dynamical equations which govern the time-evolution of both the intrinsic 3-geometry $\gamma(t)$, and its conjugate momentum field $\pi(t)$.

The Hamiltonian $H_{(N,N^i)}$ corresponds to a Hamiltonian vector field ξ_H tangential to the constraint submanifold, and the integral curves of this vector field are the classical dynamical trajectories of the Hamiltonian.[2] The one-parameter family of diffeomorphisms determined by the Hamiltonian vector field, are so-called 'canonical transformations' of the phase space; i.e., they are symplectic diffeomorphisms.

Given any initial data $(\gamma(0),\pi(0))$ which lies in the constraint submanifold, and given a choice of (N,N^i), the 'geometrodynamical' equations will evolve the initial data into a one-parameter family of pairs $(\gamma(t),\pi(t))$ on the 3-manifold Σ, each of which satisfies the constraint equations. Given initial data which satisfies the constraint equations, the dynamical evolution remains within the constraint submanifold.

Alternatively, one can think of the dynamical equations as evolving the initial data into a 4-dimensional space-time $(\mathcal{M},g)$ which satisfies the Einstein equations, ($G_{\mu\nu}=0$ if the absence of matter fields has been assumed). The 4-dimensional space-time $(\mathcal{M},g)$ contains the initial data on an embedded hypersurface Σ_0, and the 4-manifold $\mathcal{M}$ will be diffeomorphic to $\mathbb{R}\times\Sigma$.

If one thinks not merely in terms of evolving fields $(\gamma(t),\pi(t))$ upon a 3-manifold Σ, but also in terms of the 4-dimensional space-time $(\mathcal{M},g)$ which evolves from the initial data, then both H_N and H_{N^i} correspond to one-parameter groups of diffeomorphisms of $\mathcal{M}\cong\mathbb{R}\times\Sigma$, (Baez 1996a):

1. The flow on Γ generated by H_{N^i} corresponds to the N^i-diffeomorphisms of $\mathbb{R}\times\Sigma$ which map each $\Sigma_t = t\times\Sigma$ onto itself. For each such diffeomorphism $\psi:\Sigma_t\to\Sigma_t$, the pullback $\psi^*(\gamma_t)$ is isometric with γ_t.

2. The flow on Γ generated by H_N corresponds to diffeomorphisms of $\mathcal{M}$ which map each hypersurface $t_1\times\Sigma$ forward in time onto another hypersurface, $t_2\times\Sigma$, without any diffeomorphism of Σ. Such a diffeomorphism

[2] In simple terms, for a pair of conjugate variables (q_i,p_i), the Hamiltonian flow vector field is $\xi_H=(\partial H/\partial p_i,-\partial H/\partial q_i)$. The gradient of the Hamiltonian is a vector field normal to the constraint submanifold, and one rotates the gradient clockwise by ninety degrees to get a vector field tangential to the constraint submanifold.

ϕ_t maps Σ_{t_1} onto Σ_{t_2}, and in general γ_{t_1} will not be isometric with γ_{t_2}. However, in terms of the entire space-time diffeomorphism, $\phi_t : \mathcal{M} \rightarrow \mathcal{M}$, the pullback $\phi_t^*(g)$ *is* isometric with g.

One of the unusual aspects of canonical general relativity is that different choices of the lapse and shift (N, N^i) result in different Hamiltonians upon the phase space. If one fixes an initial data-point $(\gamma(0), \pi(0))$ in the constraint submanifold, but one varies the lapse and shift, then one will generate different time evolution curves in phase space, passing through $(\gamma(0), \pi(0))$. It is only when one fixes (N, N^i), when the Hamiltonian is fixed, that there is a unique dynamically possible history which passes through the point $(\gamma(0), \pi(0))$.

The Lorentzian 4-manifolds which are generated from $(\gamma(0), \pi(0))$ by different lapse and shift, will not necessarily be the same, but they will be isometric. i.e. they will be related to each other by space-time diffeomorphisms. Hence, one can say that a *physically* unique space-time is generated by admissible initial data. Given the initial data $(\gamma(0), \pi(0))$, a diffeomorphism equivalence class of Lorentzian 4-manifolds is uniquely determined. The lapse and shift one specified to evolve the initial data will define a preferential foliation of the 4-manifold. i.e. one will have a preferential diffeomorphism between $\mathcal{M}$ and $\mathbb{R} \times \Sigma$.

At this juncture, a crucial fact about the canonical formulation arises, which will be vital to assessing the viability of Loop Quantum Gravity as a physical theory: the selection of a possibly time-dependent lapse and shift field expresses the many-fingered nature of time in general relativity. At each spatial point $x \in \Sigma_t$, the local flow of time is specified by the lapse and shift vector fields.

Given a default foliation $\mathcal{M} \cong \mathbb{R} \times \Sigma$, one can define a timelike vector field ξ on $\mathcal{M}$ consisting of forward-pointing unit normals on each hypersurfaces Σ_t. Each choice of lapse N and shift N^i then defines a vector field $\zeta = N\xi + \vec{N}$. The integral curves of this vector field represent the chosen time evolution in $(\mathcal{M}, g)$. The flow of ζ defines a one-parameter family of diffeomorphisms of $\mathcal{M}$. In effect, the vector field $\zeta = N\xi + \vec{N}$ decomposes the chosen time evolution into a component parallel to the Σ_t-unit normals and a component perpendicular to those normals. Hence, the 'flow of time' is a combination of a diffeomorphism flow in Σ and a flow normal to the Σ_t-hypersurfaces in $\mathcal{M}$.

At each point $p \in \mathcal{M}$, the normalized timelike vector field $U = \zeta(p)/||\zeta(p)||$ effectively specifies an observer field[3] across space, and thereby determines a local time standard, defined by the proper time of the integral curves of U. Moreover, the observer field also defines a local rest space $U^{\perp}$, a spacelike subspace of the tangent space. Whilst the symmetry group of the tangent space $T_p\mathcal{M}$ is some component(s) of the Poincare group or the Lorentz group, (possibly $SO(3,1)$ or $SO_0(3,1)$), the symmetry group of the local rest space is $SO(3)$, the group of rotations in 3-dimensional Euclidean space.

We will see below that the proponents of LQG refer to the selection of a foliation as a 'choice of time gauge', and consider the selection of a time gauge to reduce the symmetry group from $SO(3,1)$ to $SO(3)$. However, note carefully

[3] i.e., a timelike, future-pointing unit vector field.

that whilst the symmetry group of each local rest space is indeed $SO(3)$, local Lorentz covariance still applies. The local space-time symmetry group is still a component(s) of the Poincare group or the Lorentz group, such as $SO_0(3,1)$. In particular, the local symmetry group still includes the 'boosts' between observers in local relative motion, not just the local spatial rotations.

To spell this out, consider the case of a universe which can be foliated by a one-parameter family of homogeneous and isotropic spacelike hypersurfaces, as represented by the Friedmann-Robertson-Walker (FRW) family of cosmological models. There are many foliations of each such space-time, but only one in which the embedded Σ_t are homogeneous and isotropic. There is an observer field U in such a model, defined by the future-pointing timelike unit vector field which is normal to each Σ_t. This observer field defines a local rest-space $U^\perp$, and the proper time of the integral curves of U defines a local time-standard. U-observers in such a FRW universe would detect the Cosmic Microwave Background Radiation (CMBR) to be isotropic. i.e., they would measure the temperature of its black-body spectrum to be the same in every direction.

Having defined this foliation, along with its local rest-spaces and local time-standards, one can calculate the measurements of observers who are in motion with respect to the CMBR. For example, because the Earth belongs to a galaxy which itself belongs to a galaxy cluster that possesses a peculiar velocity with respect to this local rest-space, we detect the CMBR to be redshifted in one particular direction on our celestial sphere, and blue-shifted in the opposite direction. We can use the boosts which transform between observers in relative local motion, travelling at different relative velocities, to calculate the redshifts and blueshifts measured by different local observers.

Observers in a different state of motion at the same point p, will be represented by different timelike vectors at p. Two distinct timelike vectors $V, W \in T_p\mathcal{M}$ will have different local rest spaces, $V^\perp$ and $W^\perp$, and different celestial spheres.[4] This results in the aberration of light: different observers will disagree about the position of a light source, (Sachs and Wu 1977, p46). Moreover, different observers at a point p will measure the same photon of light to possess different energies. Where Y denotes the energy-momentum tangent vector of a null geodesic at p, representing a photon, the observer (p, V) would measure it to possess energy $e = -g(Y, V)$, whilst (p, W) would measure its energy to be $e = -g(Y, W)$.

The observers V and W are related, not by an element of $SO(3)$, but by an element of the Lorentz group $SO_0(3,1)$.

So part of the routine use of relativity, as a physical theory, involves both the empirical definition of a global foliation, using the CMBR, and consistent with that, the application of the Lorentz group to transform between observers in local relative motion. This requirement for 'local Lorentz covariance' is distinct from general covariance (i.e, space-time diffeomorphism invariance), and is a critical part of the physical content of relativity. We shall return to this matter

[4]The observer's celestial sphere is the sphere of unit radius in their local rest-space.

in a later section.

The canonical formulation of general relativity also exposes another notorious interpretational problem, which is inherited by LQG: the problem of time.

Recall that time evolution in classical canonical general relativity is represented by a one-parameter family of canonical transformations of the phase space. This family is generated by a function on phase space, the Hamiltonian, which vanishes on the constraint submanifold. Such a family of canonical transformations is normally considered to be a family of gauge transformations, (Ashtekar and Geroch 1974, p1220). Thus, it is claimed, the time-evolution of a universe is represented by a succession of gauge transformations.

Furthermore, gauge transformations are generally considered to link physically equivalent states; gauge equivalence classes eliminate redundancy in the mathematical description of physical systems. Hence, it might be argued, if time is represented as a succession of gauge transformations, then time must be physically redundant.

Possible responses to this claim include:

1. These are not gauge transformations.

2. There are different types of gauge transformations, some of which involve physical redundancy and some of which don't. These belong to the types which don't.

3. The same group of transformations can possess different, but related group-actions. One action might define a group of gauge transformations, whilst another doesn't.

The third response may be appropriate in the case of canonical general relativity. The root of the problem lies in the dual role played by the space-time diffeomorphisms. Under one action, these are used to sweep a 3-dimensional spacelike hypersurface Σ through a 4-dimensional space-time, thereby evolving the metric and the matter fields on the pullback Σ; the 3-dimensional configuration evolves from one physically different configuration to another. But the space-time diffeomorphisms also act on the 4-dimensional manifold $\mathcal{M}$, and this action simply transforms one physically equivalent history to another. Thus, in their first role, the space-time diffeomorphisms are *not* gauge transformations, but in their second role, they are.

Returning to the main expository thread of this section, we now reach the point of canonical quantization. Using the 'old' variables, (γ_{ij}, π^{kl}), the need to satisfy the canonical commutation relations (CCRs) motivated the following choice of operators to represent the canonical variables:

$$\hat{\gamma}_{ij}(x)\Psi(\gamma) = \gamma_{ij}(x)\Psi(\gamma)$$

$$\hat{\pi}^{kl}(x)\Psi(\gamma) = -i\hbar\frac{\delta\Psi(\gamma)}{\delta\gamma_{kl}(x)}$$

for $x \in \Sigma;\ i, j = 1, ..., 3$, and a notional state-function $\Psi(\gamma)$ defined on the configuration space of 3-dimensional metrics.

Already, concerns arise about the meaning of these expressions. For example, as Prugovecki points out, (1992, p348), this choice of operators assumes that functional differentiation can be treated in the same way as function differentiation.

The Dirac quantization programme for a constrained system, as applied to general relativity, then proceeds as follows: The classical constraint variables are functions of the canonical variables, so formal expressions for these operators are obtained by substituting the operators representing the canonical variables into the classical functions:

$$\hat{H}_0(x; \hat{\gamma}_{ij}(x), \hat{\pi}^{kl}(x)]$$

$$\hat{H}_i(x; \hat{\gamma}_{ij}(x), \hat{\pi}^{kl}(x)]$$

for $x \in \Sigma;\ i, j = 1, ..., 3$.

With constraint operators formally defined, the next step of the Dirac quantization method is to formulate quantum versions of the classical constraint equations. Whilst the classical constraints are conditions which must be satisfied by classical states (γ, π) if those states are to be physically relevant, the quantum constraints are conditions which must be satisfied by a candidate quantum state vector Ψ if that state is to be physically relevant. The quantum constraint equations on physically admissible state functions Ψ take the form:

$$\hat{H}_0(x; \hat{\gamma}_{ij}(x), \hat{\pi}^{kl}(x)]\Psi = 0$$

$$\hat{H}_i(x; \hat{\gamma}_{ij}(x), \hat{\pi}^{kl}(x)]\Psi = 0$$

for $x \in \Sigma;\ i, j = 1, ..., 3$

Hence, physical states must lie in the kernel of all the constraint operators. Typically, it is proposed that there is a provisional vector space V, upon which the canonical operators and constraint operators have been defined, which contains a subspace V_{phys} of physical states, the kernel of all the constraint operators. In particular, state-functions $\Psi(\gamma)$ which satisfy the quantum momentum constraint equations would have the same value on isometric 3-geometries.

The quantum constraint equations are time-independent, so the problem of time rears it head again. Attempts to solve the problem have generally invoked the notion of 'intrinsic' time. The idea here is that time can be found in the domain of the wave-function. It is proposed that the degrees of freedom can be split into those which are 'physical', and those which are 'non-physical'. The physical degrees of freedom are sufficient to pin down the configuration, whilst the non-physical are redundant degrees of freedom, which purportedly contain information about intrinsic time.

The internal time is treated as a function $T[\gamma, \phi]$ of the configuration, (where ϕ denotes the matter field), and $\Psi[\gamma, \phi]$ gives the probability amplitude of the physical configuration $(\gamma, \phi)_{phys}$ at the internal time $T[\gamma, \phi]$. By allowing the internal time to vary, one would have a varying probability distribution over the possible physical configurations, so one could write the state-function as

$$\Psi[\gamma, \phi] = \Psi[(\gamma, \phi)_{phys}, T] = \Psi_T[\gamma, \phi]_{phys} .$$

Unfortunately, viable candidates for such internal time variables are difficult to find, and whilst the problem of time is a problem for Loop Quantum Gravity as much as it was a problem for metric-based formulations, it is not the focus of this paper.

So concludes our review of metric-based canonical quantization. Despite finding some application in 'mini-superspaces', where the configuration space variables could be reduced to a finite number, the metric-based approach to the canonical quantization of gravity foundered due to the intractability of its mathematics. In the late 1980s and early 1990s, however, a new hope arose, based upon a different choice of canonical variables. It is to these that we now turn.

3 Ashtekar's new variables

Using Ashtekar's 'new variables', canonical general relativity can be cast in the form of a canonical gauge theory, (Ashtekar 1986). In the canonical formulation, Yang-Mills gauge theories possess a so-called 'Gauss constraint'. Formulated as a gauge theory with Ashtekar's new variables, canonical general relativity can be mathematically treated as a gauge theory, albeit one with additional constraints to the Gauss constraint.

To obtain a gauge theory, Ashtekar first complexified general relativity. Given the specification of a Lorentzian space-time metric g on a 4-dimensional manifold $\mathcal{M}$, the bundle of oriented orthonormal frames can be equipped with a real $SO(3,1)$-connection, i.e., a real $\mathfrak{so}(3,1)$-valued 1-form.

$SO(3,1)$ is often referred to in this context as the Lorentz group. Strictly speaking, however, the Lorentz group is $O(3,1)$, a disconnected group which possesses four components, one of which contains the isometry that reverses the direction of time, another of which contains the isometry that performs a spatial reflection, and another of which contains the isometry that both reverses the direction of time and performs a spatial reflection. The identity component $SO_0(3,1)$ preserves both the direction of time and spatial parity. The identity component of the Lorentz group, $SO_0(3,1)$, is variously referred to as the *restricted* Lorentz group, or the *proper isochronous* Lorentz group. Note that the universal cover of $SO_0(3,1)$ is $SL(2,\mathbb{C})$, a fact which will be important below.

Ashtekar complexified the real $SO(3,1)$-connection into an $\mathfrak{so}(3,1)_{\mathbb{C}}$-valued 1-form. Now, the following Lie algebra isomorphisms exist:

$$\mathfrak{so}(3,1)_{\mathbb{C}} \cong \mathfrak{so}(4)_{\mathbb{C}} \cong \mathfrak{so}(3)_{\mathbb{C}} \oplus \mathfrak{so}(3)_{\mathbb{C}} .$$

The two copies of $\mathfrak{so}(3)_{\mathbb{C}}$ are treated as 'left-handed' and 'right-handed' parts of such a complex connection form. Ashtekar took the right-handed part to get an $\mathfrak{so}(3)_{\mathbb{C}}$-valued 1-form. Hence, for Ashtekar and Isham (1992), the "central ingredient" in the formalism is a complex $SO(3)$-connection.

The universal cover of $SO(3)$ is $SU(2)$, hence these Lie groups share the same Lie algebra, $\mathfrak{so}(3) \cong \mathfrak{su}(2)$. It also follows that the respective complexifications of these Lie algebras are isomorphic, $\mathfrak{so}(3)_{\mathbb{C}} \cong \mathfrak{su}(2)_{\mathbb{C}}$. Hence, a complex $SO(3)$-connection is also an $\mathfrak{su}(2)_{\mathbb{C}}$-valued 1-form.

By virtue of being complex Lie algebras, $\mathfrak{so}(3)_{\mathbb{C}}$ and $\mathfrak{su}(2)_{\mathbb{C}}$ are each vector spaces over the field of complex numbers. One also has an isomorphism between the 3-dimensional complex Lie algebra $\mathfrak{su}(2)_{\mathbb{C}}$, and the 6-dimensional real Lie algebra $\mathfrak{sl}(2,\mathbb{C})$. Thus, Ashtekar's complex $SO(3)$-connection is also isomorphic to a real $SL(2,\mathbb{C})$ connection. Coincidentally, recall that $SL(2,\mathbb{C})$ is the universal cover of the *restricted* Lorentz group, $SO_0(3,1)$. Hence, Ashtekar's new variables seem to maintain a grip on the restricted Lorentz group.

In terms of fibre bundles, Ashtekar's approach initially yields a principal bundle P with structure group $SO(4)_{\mathbb{C}}$. One takes the double cover to obtain a principal bundle $\tilde{P}$ which decomposes into a direct sum, $P_+ \oplus P_-$, each summand of which possesses structure group $SL(2,\mathbb{C})$. In the canonical version of this formulation, the configuration space for a 3-manifold $\Sigma_t \subset \mathcal{M}$, embedded in a space-time $\mathcal{M} \cong \mathbb{R} \times \Sigma$, becomes the space of connections $\mathcal{A}_+$ on P_+ *restricted to* Σ_t, (Baez 1996a).

The familiar canonical variables (γ_{ij}, π^{kl}) are replaced in Ashtekar's approach by a complex Lie algebra-valued connection form A^i_a, and a complex densitized triad field E^a_i. Assuming the choice of a trivialization of the bundle, which enables the connection on the total space of the principal bundle to be pulled down onto the base-space Σ, the a index here is a spatial index, while the i index is a Lie algebra index. So, given a coordinate chart x on Σ, and a basis X_i of the Lie algebra, one can write $A = A^i_a \, dx^a \otimes X_i$.[5]

In Ashtekar's approach we can talk interchangeably about a complex $SO(3)$-connection and a real $SL(2,\mathbb{C})$ connection. Yet Loop Quantum Gravity typically uses a real $SU(2)$ connection, so how do we get from a real $SL(2,\mathbb{C})$ connection to a real $SU(2)$ connection?

Baez (1996a), for example, states that $SL(2,\mathbb{C})$ is a complex Lie group, and "using the fact that $SL(2,\mathbb{C})$ has $SU(2)$ as a real form, one expects the kinematical state space [of quantum gravity] to be isomorphic to $L^2(\mathcal{A})$, where $\mathcal{A}$ is the space of connections on a certain $SU(2)$ bundle over [Σ]." Unfortunately, this is misleading.

To emphasize, Ashtekar's formulation involved the use of a *complex* $SO(3)_{\mathbb{C}}$

[5] Triad fields may be unfamiliar to a philosophical audience. To gain a 'feel' for what they are, note that in conventional general relativity, a real co-triad field provides the transformation which maps the components of the metric tensor into orthonormal form. If one starts with an arbitrary coordinate chart x, it determines basis vectors dx^a of the cotangent vector space at each point, and determines the expression of the spatial metric as $\gamma = \gamma_{ab} dx^a \otimes dx^b$. The co-triad fields $e^i_a(x)$ define a linear transformation at each point of space which changes the basis of each cotangent space, $\theta^i = e^i_a(x) dx^a$, so that in this new basis $\gamma = \delta_{ij} \theta^i \otimes \theta^j$. A triad field $e^a_i(x)$ is simply the corresponding transformation of the tangent vector space at each point. It is 'densitized' by bolting the square-root of the determinant of the metric onto it, $E^a_i(x) = \sqrt{\det \gamma}\, e^a_i(x)$.

connection, interchangeable with a *real* $SL(2,\mathbb{C})$ connection. In this context, $SL(2,\mathbb{C})$ is treated as a 6-dimensional real manifold, with coordinate charts mapping points of the manifold into $\mathbb{R}^6$, and a 6-dimensional real vector space as its Lie algebra. In contrast, $SU(2)$ is a 3-dimensional real manifold, equipped with a 3-dimensional real vector space as its Lie algebra. The complexification $\mathfrak{su}(2)_{\mathbb{C}}$ transforms it into a Lie algebra with 3 complex dimensions, and 3 complex dimensions enable it to be mapped to the 6 real dimensions of $\mathfrak{sl}(2,\mathbb{C})$.

It is in this sense that Livine (2009) is justified in stating that under Ashtekar's formulation, the theory is "invariant under the Lorentz group $SL(2,\mathbb{C})$ (seen as the complexified $SU(2)$ group)", and it is in this sense that Samuel (2000) can claim that Ashtekar's original formulation used a "complex $SU(2)$ connection".

The reason that Loop Quantum Gravity uses a real $SU(2)$ connection is that Barbero wrote an influential paper in 1994 which modified Ashtekar's formulation. This modified version replaces the complex $SU(2)$ connection with a real $SU(2)$ connection.

The complex $SU(2)$ connection used by Ashtekar could be written as follows:

$$A^i_a(x) = \Gamma^i_a(x) + iK^i_a(x) \ .$$

$\Gamma^i_a(x)$ is a spin connection compatible with the densitized triad, and $K^i_a(x)$ is a triadic form of the extrinsic curvature.

In contrast, Barbero's connection is a real $SU(2)$ connection of the form (Barbero 1995):

$$A^i_a(x) = \Gamma^i_a(x) + K^i_a(x) \ .$$

Barbero also replaced the complex densitized triad field in Ashtekar's formulation with a real densitized triad field, to obtain a new set of canonical variables.

Immirzi pointed out in 1995 that Barbero's connection could be generalised to (Immirzi 1997):

$$A^i_a(x) = \Gamma^i_a(x) + \beta K^i_a(x) \ ,$$

where β is any real number. The β parameter is referred to as the Immirzi parameter. Ashtekar's formulation can be recovered by rejecting the notion that β is a real number, and setting $\beta = \pm i$.

As Samuel notes, "Barbero's Hamiltonian formulation (BHF) has gained wide acceptance and is currently the basis of Loop Quantum Gravity. A lot of work has been done on the space of real $SU(2)$ connections on manifolds. Since $SU(2)$ is a compact group it has an invariant (Haar) measure, and one is able to achieve a high degree of mathematical control over the space of connections."

Indeed, as Simon remarks, "$SU(2)$ is the simplest of simple Lie groups. It is the unique rank 1, compact, connected, simply connected, semisimple Lie group," (1996, p174), and its use was responsible for much of the progress made in Loop Quantum Gravity, as we shall see in the next section.

4 Loop Quantum Gravity

We turn now to the advances made by Loop Quantum Gravity in the mid-1990s, which ultimately yielded a well-defined kinematical state space, and well-defined operators which purportedly represent spatial areas and volumes.[6]

Let's start with $\mathcal{A}$, the space of smooth real $SU(2)$ connections over Σ. Whilst this serves as the classical configuration space, the quantum configuration space is an extension $\bar{\mathcal{A}}$ of this. The extension was constructed by first noting that each smooth connection A defines a holonomy along an oriented path $\alpha : [0,1] \to \Sigma$. i.e., for a principal $SU(2)$ bundle P over Σ, each connection A defines a parallel transport map between the start-point and end-point of the path:

$$h_\alpha(A) : P_{\alpha(0)} \to P_{\alpha(1)} = \mathscr{P} \exp - \int_\alpha A \ ,$$

where $\mathscr{P}$ indicates the integral is 'path-ordered'. Using a trivialization of P, one can think of the connection as assigning an element of the Lie algebra $\mathfrak{su}(2)$ to each point of the path; the Lie algebra elements are summed, and then mapped to an element of the Lie Group $SU(2)$ by the exponential map.

One can reverse the path-integral by swapping the start-point and end-point, and one can compose two paths by setting the start-point of the second to the end-point of the first. This enables one to generalise from the space of smooth connections. A general connection $\bar{A}$ is a map on the space of such paths in Σ, which assigns an element $\bar{A}(\alpha) \in SU(2)$ to each α, such that (i) $\bar{A}(\alpha^{-1}) = (\bar{A}(\alpha))^{-1}$ and (ii) $\bar{A}(\alpha_2 \circ \alpha_1) = \bar{A}(\alpha_2) \cdot \bar{A}(\alpha_1)$, (Ashtekar 1998).

Whilst a smooth gauge transformation assigns an element of $SU(2)$ to each element of Σ, a generalised gauge transformation g acts on the start-points and end-points of paths so that $\bar{A}(\alpha) \to g(\alpha(1))^{-1} \cdot \bar{A}(\alpha) \cdot g(\alpha(0))$. The space of generalized gauge transformations is denoted as $\bar{\mathcal{G}}$, and the gauge invariant quantum configuration space is the quotient $\bar{\mathcal{A}}/\bar{\mathcal{G}}$.

Whilst this is an extremely large space, it can be treated as a 'projective limit' of a family of compact, smooth, finite-dimensional configuration spaces, each of which is labelled by a graph γ in Σ. Each graph is a collection of N edges and V vertices. The space of connections over the graph, $\mathcal{A}_\gamma$, consists of mappings which assign an element of $SU(2)$ to each edge; the space of gauge transformations over the graph $\mathcal{G}_\gamma$ consist of mappings which assign an element of $SU(2)$ to each vertex of the graph. If we denote an edge by e, its start-point by v_{e-}, and its end-point by v_{e+}, then $\mathcal{G}_\gamma$ has the following action on $\mathcal{A}_\gamma$: $A_\gamma(e) \to g(v_{e+})^{-1} \cdot A_\gamma(e) \cdot g(v_{e-})$.

The gauge invariant configuration space associated with the graph γ is simply $\mathcal{A}_\gamma/\mathcal{G}_\gamma$. $\mathcal{A}_\gamma$ is isomorphic with $[SU(2)]^N$, and $\mathcal{G}_\gamma$ is isomorphic with $[SU(2)]^V$, hence these are compact and finite-dimensional spaces. The projective limit $\bar{\mathcal{A}}/\bar{\mathcal{G}}$ is also compact, and admits a regular Borel measure which enables the construction of a Hilbert space $L^2(\bar{\mathcal{A}}/\bar{\mathcal{G}})$ of square-integrable functions. In fact,

[6] Here the exposition closely follows Ashtekar (1998).

the natural measure on $\bar{\mathcal{A}}$ is induced by the Haar measure on $SU(2)$, and is diffeomorphism invariant.

The Hilbert space $\mathscr{H} = L^2(\bar{\mathcal{A}}/\bar{\mathcal{G}})$ is the kinematic state space of the quantum theory. It is infinite-dimensional. Moreover, for each graph γ, there is a subspace of states $\mathscr{H}_\gamma = L^2(\bar{\mathcal{A}}_\gamma/\bar{\mathcal{G}}_\gamma)$ defined upon the graph, which is itself an infinite-dimensional subspace. It consists of functions of the form $\Psi_\gamma(A_\gamma) = \psi(A_\gamma(e_1), \ldots A_\gamma(e_N))$, for smooth functions ψ on $[SU(2)]^N$, where $e_1 \ldots e_N$ are edges of the graph. Ashtekar refers to these states as "1-dimensional polymer-like excitations of geometry/gravity," (1998, p181). In a similar vein, Ashtekar *et al* assert that "elements of these $[\mathscr{H}_\gamma]$ describe elementary quanta of geometry; to obtain classical geometries one needs to coherently superpose a large number of them," (Ashtekar, Rovelli and Reuter, 2014, p20).

However, the kinematic Hilbert space can also be decomposed into an infinite direct sum of *finite*-dimensional, orthogonal subspaces, (Ashtekar, Rovelli, and Reuter 2014, p18):

$$\mathscr{H} = \oplus_{\gamma,\vec{j}} \mathscr{H}_{\gamma,\vec{j}} \ .$$

Each subspace is labelled by not just a graph γ, but by the assignment of an irreducible representation of $SU(2)$ to each edge of the graph.[7] The representations of $SU(2)$ are parameterized by half-integers j, hence the assignment of such a representation to each edge can be denoted by a vector of such half-integers $\vec{j}$. Such a structure is called a 'spin network', and elements of $\mathscr{H}_{\gamma,\vec{j}}$ are called spin network states.

The spin networks embedded in the 3-manifold Σ can be used to define functionals $\Psi_{\gamma,\vec{j},\vec{i}}$, called spin-network states, on the space of $SU(2)$-connections. In essence, $\Psi_{\gamma,\vec{j},\vec{i}}$ can be constructed as follows: take the value of the connection along each edge; use the edge-dependent representations $\vec{j}$ to obtain operators along each edge, and form the tensor product of all those operators; select $\vec{i}$, the assignment of an operator at each vertex which intertwines between the representations associated with the edges coming into and going out of that vertex; tensor all the edge and vertex operators together; then contract to obtain a number. This is the value $\Psi_{\gamma,\vec{j},\vec{i}}(A)$ assigns to the connection A, (Baez 1996a; Norton 2016).

In more detail, the procedure is as follows: Denote the representation assigned to edge e as ρ_e. i.e., ρ_e is a homomorphism $\rho_e : SU(2) \to \mathrm{End}(V_e)$, where $\mathrm{End}(V_e)$ is the group of endomorphisms of some vector space V_e. Each edge e in γ maps $A \in \mathcal{A}_\gamma$ to the linear operator $L_e(A) = \rho_e(A_e)$. A linear operator can

[7] The irreducible representations of $SU(2)$ are parameterized by $j \in \frac{1}{2}\mathbb{Z}_+$. As one possible realisation of these representations, let V_j, for all $j \in \frac{1}{2}\mathbb{Z}_+$, denote the space of homogeneous polynomials of degree j on $\mathbb{C}^2$, (Sternberg 1994, p181). i.e., V_j is the space of functions

$$p(z_1, z_2) = \sum_{k=0}^{2j} a_k z_1^{2j-k} z_2^k \ .$$

$V_{\frac{1}{2}} = \mathbb{C}^2$, the standard representation of $SU(2)$ on $\mathbb{C}^2$, and this determines a representation of $SU(2)$ upon each of the function spaces.

be written as a tensor with one contravariant and one covariant index, $L_e(A)^i_k$. Take the tensor product $L = \bigotimes_e L_e(A)$. In terms of indices, this is a tensor $L_e(A)^{i_1,\ldots,i_N}_{k_1,\ldots,k_N}$. Next, for each vertex $v \in \gamma$, form the set of edges $S(v)$ which have v as a start-point, and the set of edges $T(v)$ which have v as an end-point, (Baez 1996a). Next, define an 'intertwining' operator:

$$I_v = \bigotimes_{e \in T(v)} V_e \rightarrow \bigotimes_{e \in S(v)} V_e \ .$$

In other words, I_v is a linear endomorphism between the vector spaces which the representations act upon. Each I_v can also be written as a tensor with contravariant and covariant indices, $(I_v)^{l_1,\cdots}_{m_1,\ldots}$, with one contravariant index for each edge $e \in S(v)$, and one covariant index for each edge $e \in T(v)$.[8] Next form the tensor product of all the intertwining operators:

$$I = \bigotimes_v I_v \ .$$

If we form the tensor product $L \otimes I$, it has the index form $(L \otimes I)^{i_1,\ldots,i_N,l_1,\ldots l_N}_{k_1,\ldots,k_N,m_1,\ldots,m_N}$. Each contravariant index of L corresponds to a covariant index of I, and vice versa, so one can contract the entire tensor to get a number (Baez 1996a). This number is simply $\Psi_{\gamma,\vec{j},\vec{i}}(A)$.

The set of all spin network states $\Psi_{\gamma,\vec{j},\vec{i}}$, taken over all graphs γ, all assignments $\vec{j}$ of representations to the edges, and all assignments of intertwining operators $\vec{i}$ to the nodes, spans the kinematic Hilbert space of gauge invariant states $\mathscr{H} = L^2(\bar{\mathcal{A}}/\bar{\mathcal{G}})$, (Baez 1996b).

For a fixed graph γ, the set of all spin network states for all assignments $\vec{j}$ of representations to the edges, and all assignments of intertwining operators $\vec{i}$ to the nodes, spans the subspace of gauge invariant states on that graph $\mathscr{H}_\gamma = L^2(\bar{\mathcal{A}}_\gamma/\bar{\mathcal{G}}_\gamma)$, (Baez 1996b).

Such functionals turn out to be eigenvectors of well-defined operators on the kinematic Hilbert space, which purportedly represent the area of surfaces, and the volume of regions in the 3-manifold. Furthermore, these operators have discrete spectra.

For a state $\Psi_\gamma(A) = \psi(A(e_1), \ldots A(e_N))$ specified by some smooth function ψ on $[SU(2)]^N$, the area operators are basically constructed from left/right invariant vector fields on $SU(2)$.

For a specified 2-dimensional surface $\mathcal{S} \subset \Sigma$, one first defines an operator X^i_I corresponding to each edge I of γ which intersects the surface $\mathcal{S}$. The i index here labels a basis τ^i of the Lie algebra $\mathfrak{su}(2)$. "At the vertex of intersection each edge contributes via the Lie derivative along the i-th right or left invariant vector field on the copy of $SU(2)$ associated with that edge (depending on whether the edge is oriented to be outgoing at the vertex or incoming," (Ashtekar 1998, p184). For example, when the edge e_I is outgoing,

$$(X^i_I \cdot \psi)(A_\gamma(e_1), \ldots A_\gamma(e_N)) = (A(e_I)\tau^i)^A_B \frac{\partial \psi}{\partial (A(e_I))^A_B} \ .$$

[8] Baez (1996a) and Norton (2016) use the opposite index convention.

The indices A, B in expressions such as $(A(e_I))^A_B$ here, are spinor indices, taking values $0, 1$, and used to identify the elements of the $SU(2)$ matrix.

To construct the area operators, identify pairs of edges (I, J) which intersect the surface, and form an operator based upon a sum of terms containing $X^i_I X^i_J$. Specifically, for each point $v \in \mathcal{S}$, one can form a vertex operator which can be regarded as a Laplacian operator on the kinematic Hilbert space:

$$\Delta_{S,v} = \sum_{I_v, J_v} \kappa(I_v, J_v) X^i_I X^i_J \ ,$$

where I_v and J_v label edges of the graph γ which have $v \in \mathcal{S}$ as an intersection vertex. There will only be a finite number of points where the graph intersects $\mathcal{S}$, and the area operator can be defined as the sum of the square-roots of the Laplacians over all such points:

$$\hat{A}_{\mathcal{S}} = \frac{\ell_P^2}{2} \sum_{v \in \mathcal{S}} \sqrt{-\Delta_{S,v}} \ ,$$

where ℓ_P is the Planck length.

The area operators are self-adjoint operators which leave the space of states $\mathscr{H}_\gamma = L^2(\mathcal{A}_\gamma/\mathcal{G}_\gamma)$ associated with each graph γ invariant. $L^2(\mathcal{A}_\gamma/\mathcal{G}_\gamma)$ is a space of functions on a compact manifold, isomorphic to $[SU(2)]^N$, and the restriction of each area operator to such a subspace is a sum of elliptic differential operators on a compact manifold, hence the spectrum of each area operator on $\mathscr{H}_\gamma$ is discrete. Moreover, the area operators can be extended to the entire kinematic state space $\mathscr{H}$, upon which the spectrum of each area operator also transpires to be discrete, (Ashtekar 1998, p188). Similar results can be obtained for the volume operators.

The simplest eigenvectors Ψ_γ of $\hat{A}_{\mathcal{S}}$ are those associated with graphs γ which possess N intersections with $\mathcal{S}$, each of which is such that there is exactly one incoming edge and one outgoing edge. In this case, Ψ_γ is an eigenvector of $\hat{A}_{\mathcal{S}}$ with eigenvalue $a_{\mathcal{S}} = N(\sqrt{3}/2)\ell_P^2$. In other words, these eigenvalues are equally spaced, with values proportional to the number of points of intersection.

The complete spectrum of the area operators can be computed, and the topology of the 2-dimensional surface $\mathcal{S}$ determines the pattern of spacing between the lowest eigenvalues. Moreover, the spacing between the area eigenvalues reduces as the area increases. (The same applies to the eigenvalues of the volume operators). "In the case of trivial $[\mathcal{S}]$ topology, for instance, there is only one non-zero eigenvalue with $a_{\mathcal{S}} < 0.5\ell_P^2$, seven with $a_{\mathcal{S}} < \ell_P^2$ and 98 eigenvalues with $a_{\mathcal{S}} < 2\ell_P^2$... The fact that the level-spacing in the spectrum of the area operator goes rapidly to zero makes it easy to visualize why the continuum picture is such an excellent approximation even on the smallest laboratory scales... probed in high energy physics," (Ashtekar 1998, p189).

5 Issues with LQG and discrete space-time

At face value, Loop Quantum Gravity has delivered some impressive results. However, the area and volume operators are only defined on a kinematic Hilbert space. To obtain a physical Hilbert space $\mathscr{H}^{phys}$ requires solution of the Gauss constraint, the spatial diffeomorphism constraint, and the Hamiltonian constraint, followed by the construction of an inner-product on the resulting space. Unfortunately, whilst the first two constraints were solvable,[9] the Hamiltonian constraint proved to be intransigent.

This failure to solve the dynamics in the canonical formulation spurred the Loop Quantum Gravity community to develop a sum-over-histories approach instead. This is commonly referred to as the theory of 'spin-foams'.

Essentially a spin-foam is a higher-dimensional analogue of a spin-network. It is a 2-complex embedded in the 4-dimensional manifold $\mathcal{M}$; i.e., it consists of a collection of faces, bounded by edges which join at vertices. Significantly, however, there is a representation of the Lorentz group $SL(2,\mathbb{C})$ associated to each face. Any slice through a spin-foam, or any boundary of it, is a spin-network in which representations of $SL(2,\mathbb{C})$ rather than $SU(2)$, are assigned to each edge.

As with any sum-over-histories approach to the dynamics of a quantum theory, one calculates the transition amplitude between an initial state and a final state by summing over all the possible interpolating histories. In the case of spin-foams, the initial state and final state correspond to a pair of spin networks, and for each interpolating 2-complex one sums over all possible combinations of representations assigned to the faces.

Putting to one side all the issues with the dynamics, there is a fundamental interpretational difficulty with LQG: what is the physical significance of the representations of $SU(2)$ associated with the edges of a spin-network? And, linked with this question, what has happened to the local Lorentz covariance of general relativity? How does one recover a classical limit of LQG if there is no local Lorentz covariance? Whilst diffeomorphism invariance and gauge invariance seem to have been sacred objectives for the LQG community, local Lorentz covariance seems to have been somewhat neglected. Given that local Lorentz covariance encodes a significant part of the empirical content of general relativity, its apparent loss poses a problem for any effort to link LQG with the physical world.

This loss of local Lorentz covariance can be traced back to the adoption of the Barbero connection in 1994. Recall that in Ashtekar's original re-formulation of general relativity as a gauge field theory, the connection was an $SL(2,\mathbb{C})$ connection on space-time. $SL(2,\mathbb{C})$ is the universal cover of the restricted Lorentz group $SO_0(3,1)$, so local gauge-invariance under this group could be interpreted as local Lorentz covariance. Moreover, Ashtekar's formulation involves the restriction of a space-time connection to a spatial connection on a hypersurface

[9] The space of states which solve the Gauss constraint and the diffeomorphism constraint is spanned by states which correspond to diffeomorphism equivalence classes of spin-networks. These equivalence classes are dubbed 's-knots'.

Σ.

As Samuel (2000) points out, not only did Barbero's modification break local Lorentz covariance, but the Barbero connection over a 3-manifold Σ cannot be interpreted as the restriction of a space-time connection. Whilst the Barbero connection is invariant under spatial diffeomorphisms, it is *not* invariant under diffeomorphisms normal to the embedding of Σ in $\mathcal{M}$.

Samuel demonstrates this with a simple example. Consider a loop γ of radius R in Minkowski space-time. Consider the trace of the holonomy of a connection around that loop. Consider that fixed loop embedded in two different hypersurfaces: one is a flat hyperplane with $t = \sqrt{1+R^2}$ in conventional global Cartesian coordinates; the other is a spacelike hyperboloid defined by $t - x^2 - y^2 = 1$. Under the Barbero connection, the value of the trace of the holonomy is different for the two hypersurfaces. Hence, Barbero's connection is not the restriction of a space-time connection, and the Barbero connection cannot provide a formulation of general relativity as a space-time gauge theory.

Giulini concurs, asserting that if the spatial connection is not a space-time gauge field restricted to a spacelike hypersurface, then "the dynamics generated by the constraints does then not admit the interpretation of being induced by appropriately moving a hypersurface through a spacetime with fixed geometric structures on it," (Giulini, 2009).

Samuel anticipates a common objection: "One sometimes sees it implied that the reduction in the gauge group from $SO(3,1)$ to $SO(3)$ takes place because of our choice of the 'time gauge'. It is indeed true that once we make this gauge choice, our freedom to make additional gauge transformations is curtailed from $SO(3,1)$ to $SO(3)$. However this does not mean that the gauge group has been reduced. The pullback of the connection to a spatial slice is still an $SO(3,1)$ connection, in spite of our gauge choice." This is essentially the same point we made in the opening section concerning the continued validity of the Lorentz group to transform between the measurements of observers in local relative motion, even after the choice of a global space-time foliation.

Moreover, as Samuel adds, "If one gives up the gauge interpretation of gravity, the Immirzi parameter appears not to be fixed by theory. This would not be a problem if the parameter disappeared from all physical predictions of the theory. However, this is not the case: the Immirzi parameter does appear in the calculated value of Black Hole entropy in Loop Quantum Gravity."

He concludes: "we argue strongly for maintaining the gauge aspect of gravity in the approach to quantum gravity, even though the gauge group is noncompact and therefore not as tractable as say, $SU(2)$. It does not appear to us a strong argument to say that we study compact gauge groups because we do not know how to deal with noncompact gauge groups with mathematical rigour. The noncompactness of the gauge group appears to us an essentially physical feature of General Relativity, which is closely related to the Minkowskian signature of the spacetime metric and light cones."

Note carefully that one can agree with Samuel that the loss of the Lorentz group, and the consequent loss of local Lorentz covariance, seriously detracts from the physical/empirical content of Loop Quantum Gravity, even if one doesn't agree that the Lorentz group is actually a gauge group *per se*.

So, what was the reaction of the LQG community to the points made in Samuel's paper? Rather unfortunately, Samuel choose to refer to his own argument as providing "some criticisms of an aesthetic nature." In Thomas Thiemann's voluminous reference work on canonical quantum general relativity, he uses this to largely dismiss Samuel's objections, asserting: "the criticism by Samuel... is really just aesthetical in nature and not an obstruction to implementing spacetime covariance... one just must not commit the mistake of thinking the $SU(2)$ connection is the pullback of a spacetime connection," (2008, p122).

On the contrary, Samuel's objections are not merely aesthetic, but are crucial to the prospective physical (i.e., the empirical) content of LQG. Note also that Thiemann conflates 'spacetime covariance' (i.e., general covariance, or spacetime diffeomorphism invariance) with local Lorentz covariance, and misconstrues Samuel's objection as questioning the ability of the Barbero Hamiltonian formulation to satisfy the requirement for spacetime covariance.

In general, expressions of disquiet about the physical content of LQG seem to have been overwhelmed by the mathematical tractability of the theory with Barbero's real-valued β. As Thiemann comments, "to date a satisfactory quantum theory has been constructed only for β real (which in turn does not mean that it is impossible to do for $\beta = i$)," (ibid).

One alternative voice was provided by Etera Livine in 2006, who boldly declared "we address a fundamental issue at the root of LQG, which is necessarily related to these questions: why the $SU(2)$ gauge group of Loop Quantum Gravity? Indeed, the compactness of the $SU(2)$ gauge group is directly responsible for the discrete spectra of areas and volumes, and therefore is at the origin of most of the successes of LQG: what happens if we drop this assumption? As we will see, this leads to a theory of Covariant Loop Quantum Gravity, which uses the same techniques and tools as LQG but whose gauge group is the Lorentz group $SL(2, \mathbb{C})$ instead of $SU(2)$."

Livine proposed a theory which doesn't suffer from the Immirzi ambiguity, which uses spin networks that assign representations of the Lorentz group to each edge, and which duly yields operators possessing a continuous area spectrum. For Livine, this approach resolves the conflict of the conventional LQG kinematics with the dynamics of spin-foams, where the kinematic spin-networks are $SU(2)$-networks, but the spin-foam networks are Lorentz group networks.

Rovelli claims that "in the context of LQG... discreteness is not imposed or postulated. Rather, it is a direct consequence of a straightforward quantization of GR," (2004, p250). Similarly, Smolin (2009) claims that "discreteness is a true generic consequence" of a class of theories containing Loop Quantum Gravity. Yet Livine's work, (ironically published in the same volume containing Smolin's paper), demonstrates very clearly that operators with a discrete spectrum are not the inevitable outcome of quantizing general relativity with the use of spin networks.

In fact, it's very clear that most practitioners of quantum gravity have a strong preconception that the quantization of space-time should yield some type of discrete structure. Ashtekar enunciates a particularly clear example of this:

"What do we mean by 'quantization' [in non-relativistic quantum mechanics]? We mean that there exist physical quantities which can take on continuous values classically but are such that the corresponding quantum operators have a discrete spectrum. For example, this is the sense in which the energy and angular momentum of the hydrogen atom are quantized. The question therefore is whether there exist geometrical quantities for which a similar quantization occurs. In differential geometry, lengths of curves, areas of surfaces and volumes of regions can take on continuous values... The question then is: Can one construct corresponding self-adjoint operators in the quantum theory and, if so, do they have discrete spectra? If so, we will say that geometry is quantized." (1998, p176).

Applying this logic, Ashtekar would reject Livine's approach to quantum gravity out-of-hand on the basis that it yields operators with continuous spectra. However, carrying such baggage into a research programme may be damaging. For a start, consider Ashtekar's claim that the energy of the hydrogen atom is quantized. This is only true for the bound system. Once ionized, the free electron can possess a continuous range of energies. Hence, the energy spectrum of the hydrogen atom is part discrete, and part continuous. This is despite the fact that the energy of a free system is represented by an elliptic differential operator, the Laplacian.

Second, consider Ashtekar's claim that in quantum theory, "there exist physical quantities which can take on continuous values classically but are such that the corresponding quantum operators have a discrete spectrum." The two most fundamental quantities in non-relativistic quantum mechanics, (in fact, the canonical variables in a Hamiltonian formulation), are position and momentum. Both of these quantities are represented in quantum theory by self-adjoint operators with a fully continuous spectrum.

The angular momentum operators, of course, *do* possess a discrete spectrum, and here another subtlety of quantum mechanics becomes apparent. Angular momentum is a function of the canonical variables, position and linear momentum, $\mathbf{l} = F(\mathbf{r}, \mathbf{p}) = \mathbf{r} \times \mathbf{p}$. To construct the operator representing any such function, one substitutes the operators representing $\mathbf{r}$ and $\mathbf{p}$ into the functional expression F. Thus, $\hat{\mathbf{l}} = F(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \mathbf{r} \times (-i\hbar\nabla)$.

However, despite the fact that $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ both possess continuous spectra, the cross-product doesn't. When the cross-product is expressed in spherical polar coordinates, and the quantum operators for position and momentum are substituted in, the result is that each component of angular momentum loses any radial dependence:

$$\hat{\mathbf{l}}_x = i\hbar(\sin\phi \frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi})$$

$$\hat{\mathbf{l}}_x = i\hbar(-\cos\phi \frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi})$$

$$\hat{\mathbf{l}}_x = -i\hbar\frac{\partial}{\partial\phi}$$

As a result of this purely angular dependence, the angular momentum operators act on functions defined on a compact space, the 2-sphere. When the square of the total angular momentum is taken, it then defines a Laplacian on the 2-sphere:

$$\hat{\mathbf{l}}^2 = \hat{\mathbf{l}}_x^2 + \hat{\mathbf{l}}_y^2 + \hat{\mathbf{l}}_z^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} (\sin\theta \frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right] .$$

As a consequence of defining an elliptic differential operator on a compact space, this operator possesses a purely discrete spectrum. Its eigenfunctions are the spherical harmonics $Y_l^m(\theta, \phi)$:

$$\hat{\mathbf{l}}^2 Y_l^m(\theta, \phi) = \hbar^2 \lambda Y_l^m(\theta, \phi) = \hbar^2 \ell(\ell + 1) Y_l^m(\theta, \phi) .$$

For a fixed $\ell = 0, 1, 2, \ldots$, the eigenvalue is $\hbar^2 \lambda = \ell(\ell{+}1)$. (These eigenfunctions are degenerate, with $m = \ell, \ell - 1, \ldots, -\ell$).

So, functional analysis is perfectly capable of generating observables with a discrete spectrum from observables with a continuous spectrum. There are two lessons to absorb here:

1. Quantization is capable of generating observables with a purely continuous spectra, observables with a purely discrete spectra, and observables with spectra that are part continuous and part discrete.

2. Quantization is capable of constructing discrete observables as functions of continuous observables.

These are not so much facts about non-relativistic quantum mechanics as facts about functional analysis. In as much as any viable theory of quantum gravity will have to be constructed with the tools of functional analysis, it follows that these facts will apply to any viable theory of quantum gravity. Even if the conviction that quantum gravity should yield some form of discrete structure is correct, it may be unnecessary to force discrete spectra into the theory at all levels. One can expect quantum gravity to be characterised, not exclusively by operators with discrete spectra, but by a mix of the discrete and the continuous.

The misconception that quantization should introduce discrete structures into a physical theory also betrays an implicit assumption that the discrete and the continuous are fundamentally different, or in opposition. The mathematical structures used in classical physics reveal this to be a false dichotomy. Consider for example the 3-dimensional differential manifolds Σ used in canonical general relativity. These are perfect exemplars of the continuum, which Loop Quantum Gravity apparently seeks to replace. Yet every 3-manifold is homeomorphic to a cell complex, an object specified in discrete terms.

Let's spell this out explicitly. An n-cell is an object which is homeomorphic with the n-ball in n-dimensional Euclidean space, $\mathbb{D}^n = \{x \in \mathbb{R}^n : \|x\| = 1\}$. For example, a 2-ball is a disc, bounded by a circle, while a 3-ball is a solid ball bounded by a 2-sphere. Any polygon is homeomorphic with a 2-ball, and is therefore a 2-cell. Any solid polyhedron is homeomorphic with a 3-ball, and is therefore a 3-cell.

A cell-complex is obtained by pasting together any number of cells, so that the faces of the cells are either disjoint, or so that they coincide completely. A 3-dimensional cell-complex is obtained by pasting together 3-cells in such a way that the faces, edges and vertices of the cells are either disjoint, or they coincide completely.

The most interesting type of cell is a simplex. A 0-simplex is a point, or 'vertex', a 1-simplex is a line segment, or 'edge', a 2-simplex is a triangle, and a 3-simplex is a solid tetrahedron. By pasting together simplices, one obtains a simplicial complex, (see Stillwell 1992, p23-24). A 3-dimensional simplicial complex is obtained by pasting together solid tetrahedra.

Moise's triangulation theorem for 3-manifolds, (Stillwell 1992, p25 and p242), demonstrates that every 3-manifold is homeomorphic with a simplicial complex; one says that every 3-manifold can be 'triangulated'. In fact, every n-manifold can be triangulated for $n \leq 3$.

The discrete and the continuous are intimately related here. This is a fact about the algebraic topology of manifolds, which is independent of any quantization technique. Yet these discrete structures are forwarded within Loop Quantum Gravity as validation of the discrete credentials of the theory. It is, for example, pointed out that every spin-network is 'topologically dual'[10] to a 3-dimensional simplicial complex, each node of the network sitting inside a tetrahedron, and each edge of the network crossing a triangular face, (Oriti 2006).

Ashtekar, Rovelli and Reuter claim that "perhaps the simplest way to visualize the elementary quanta [of geometry] is to introduce a simplicial decomposition...of the 3-manifold Σ and consider a graph γ which is dual [to the simplicial complex]: Each cell...is a topological tetrahedron T_n, dual to a node n of γ; each face F_e...is dual to a link e...consider a basis Ψ_{γ,v_n,a_e} in $\mathscr{H}_{\gamma,\vec{j}}$ that simultaneously diagonalizes the volume operator associated with the tetrahedron T_n, and the area operators associated with the faces F_e, for all n, e. Each of these spin-network states describes a specific elementary quantum geometry. One can think of the node n as a 'grain' or a 'quantum' of space captured in the (topological) tetrahedron T_n...each T_n has a well defined volume v_n and each of its faces F_e has a well-defined area a_e. But now the v_n, a_e are discrete. More importantly...T_n no longer has the sharp geometry of a geometrical tetrahedron...In particular, operators describing angles between any two distinct faces F_e, $F_{e'}$ of a T_n are not diagonal in the basis...These properties of the quantum geometry associated with the basis Ψ_{γ,v_n,a_e} are closely analogous to the properties of angular momentum captured by the basis $|\ell, m>$ in quantum mechanics: it too diagonalizes only some of the angular momentum operators, leaving values of other angular momentum observables fuzzy. Thus, each of the elementary

[10]The dual K^* of an n-dimensional simplical complex K assigns to each k-dimensional subsimplex $\sigma_k \subset K$, an $(n-k)$-dimensional subsimplex $\sigma^*_{n-k} \subset K^*$, so that the dual complex of the dual complex is the original complex, $K^{**} \cong K$, (Benedetti 2007, p15).

cells in the simplicial decomposition is now a 'tetrahedron' in the same heuristic sense that a spinning particle in quantum mechanics is a rotating body'," (2014, p20, with notation modified for consistency with this paper).

The architects, proponents and interpreters of LQG then pose themselves the task of explaining how the classical continuum can emerge from these discrete spin-network structures. Wüthrich writes:

"Essentially, each node (and only the nodes) in the network contributes a term to the sum of the volume of a region. On each node, there sits an 'atom' of space with volume v_n, as it were. These elementary grains of space are separated from each other by their surfaces of contiguity. Just as the volume operator receives contributions from the nodes of a region, the area operator acquires contributions from all the links that intersect the surface... the 'size' of the surface connecting adjacent 'chunks' of 'space' is constructed from the spin representations sitting on the relevant links. Thus, the smooth space of the classical theory is supplanted by a discrete quantum structure displaying the granular nature of space at the Planck scale. Continuous space as we find it in classical theories such as GR and as it figures in our conceptions of the world is a merely emergent phenomenon... LQG *predicts* the existence of indivisible quanta of volume, area, and length, as well as their spectra (up to a constant). Importantly, this discreteness was a *result* of the loop quantization, rather than an *assumption*." (2017, p314).

Yet the simplicial decompositions, and their dual graphs, were already present in the algebraic topology of the classical theory, and the discreteness of the LQG operator spectra follows not from the use of spin networks, but from the choice of a compact gauge group, and the choice of a compact gauge group *is* one of the assumptions of LQG.

Retain the spin networks, but replace the compact group with a non-compact group, and the grains of sand which were falling through one's fingers suddenly fuse themselves back into a glassy continuum. Hence, as Butterfield and Isham comment, in a different context, "the clash... between the disparate bases of... quantum theory and general relativity need not be so straightforward as the contradiction between discreteness and continuity," (2001, p34).

The interpretational problems posed by non-commuting observables remain if the fundamental observables in quantum gravity possess a continuous rather than a discrete spectrum, but the problem then becomes the familiar one of how the definite and local classical world can emerge from the fuzzy and potentially non-local quantum world. As Huggett and Wüthrich put it: "the actually existing and physically fundamental structure is supposed to be a quantum superposition of something like these spin networks, and not just a single spin network. Since all the different structures in the superposition will have a different connectivity (and perhaps different cardinality), and in this mathematical sense be different structures altogether, what is local in one term of the superposition will in general not be local in others.... How such local, i.e. topological, structures like relativistic spacetimes emerge from spin networks is at present little understood," (2012, p6).

6 Conclusions

In itemised form, we have arrived at the following propositions:

- Some of the discrete structures commonly associated with Loop Quantum Gravity, such as cell-complexes and graphs, are homeomorphic with, or topologically dual to the mathematical structures of classical general relativistic space-time; they are not borne of quantization alone.

- The tools of functional analysis associated with any quantum theory generally yield self-adjoint operators with continuous and discrete spectra. Moreover, observables with a discrete spectra can be constructed from observables with a continuous spectrum. The notion that quantization should yield operators with a discrete spectrum is a general misconception, and an assumption which has critically influenced the development of Loop Quantum Gravity.

- Spin networks are not necessarily associated with operators that possess a discrete spectrum. The discreteness is, instead, a consequence of assuming a compact gauge group. This is a contingent assumption: if a non-compact group is used, spin networks lead to operators with a continuous spectrum.

- The Barbero spatial connection, which underpins LQG, cannot be interpreted as the restriction of a space-time gauge field to an embedded spatial hypersurface.

- The preconception that quantization should yield operators with a discrete spectrum has led to the rejection of the Lorentz group as a gauge group, and this had divested Loop Quantum Gravity of the physical content provided by local Lorentz covariance.

- At present, it appears that the tractability of LQG is heavily dependent upon the compact nature of $SU(2)$. It also appears that LQG lacks a classical limit which reproduces local Lorentz covariance. Given the non-compact nature of the Lorentz group, this leads to the following vexatious conclusion:

  > *In as much as Loop Quantum Gravity is tractable, it is physically irrelevant, and is as much as it is physically relevant, it is intractable.*

References

[1] Ashtekar, A. and Geroch, R. (1974). *Quantum theory of gravitation*, Rep.Prog.Phys. 37, pp1211-1255.

[2] Ashtekar, A. (1986). *New Variables for Classical and Quantum Gravity*, Phys. Rev. Lett. 57 2244.

[3] Ashtekar, A. and Isham, C. (1992). *Representation of the holonomy algebras of gravity and non-abelian gauge theories*, Class. Quant. Grav. 9, pp1433-1467.

[4] Ashtekar, A. (1998). *Geometric issues in quantum gravity*, in *The geometric universe*, eds. S.A.Huggett, L.J.Mason, K.P.Tod, S.T.Tsou, N.M.J. Woodhouse, Oxford University Press, pp173-194.

[5] Ashtekar, A., Reuter, M, and Rovelli, C. (2015). *From general relativity to quantum gravity*, in *General relativity and gravitation: A centennial perspective*, eds. A.Ashtekar, B.K.Berger, J.Isenberg, and M.MacCallum. pp553-610.

[6] Baez, J.C. (1996a). *Spin networks in nonperturbative quantum gravity*, in *The interface of knots and physics*, ed. L.Kauffman, American Mathematical Society.

[7] Baez, J.C. (1996b). *Spin network states in gauge theory*, Adv. Math. 117, pp253-272.

[8] Barbero, J.F. (1995). *Real Ashtekar Variables for Lorentzian Signature Space-times*, Phys. Rev. D., 51, pp5507-5510.

[9] Benedetti, D. (2007). *Quantum Gravity from Simplices: Analytical Investigations of Causal Dynamical Triangulations*, https://arxiv.org/pdf/0707.3070.pdf

[10] Butterfield, J. and Isham, C.J. (2001). *Spacetime and the Philosophical Challenge of Quantum Gravity*, in *Physics meets Philosophy at the Planck Scale*, ed. C. Callender and N. Huggett, Cambridge University Press, pp33-89.

[11] Giulini, D.(2009). *The superspace of geometrodynamics*, Gen. Relativ. Gravit. 41(4), pp785-815.

[12] Huggett, N. and Wüthrich, C. (2013). *Emergent spacetime and empirical (in)coherence*, Studies in the History and Philosophy of Modern Physics 44, pp276-285.

[13] Immirzi, G. (1997). *Real and complex connections for canonical gravity*, Class. Quant. Gravity 14, L177-181.

[14] Livine, E.R. (2009). *Towards a Covariant Loop Quantum Gravity*, in *Approaches to quantum gravity*, ed. D.Oriti, Cambridge University Press, pp253-270.

[15] Norton, J. (2015). *Loop quantum ontology: spin-networks and space-time*, preprint, http://philsci-archive.pitt.edu/12016/

[16] Prugovecki, E. (1992). *Quantum geometry*, Kluwer.

[17] Rovelli, C. (2004). *Quantum Gravity*, Cambridge University Press.

[18] Sachs, R.K. and Wu, H. (1977). *General relativity for mathematicians*, Springer-Verlag.

[19] Samuel, J. (2000). *Is Barbero's Hamiltonian formulation a Gauge Theory of Lorentzian Gravity?*, Class.Quant.Grav.17:L141-L148.

[20] Simon, B. (1996). *Representations of Finite and Compact Groups*, American Mathematical Society, Graduate Studies in Mathematics, Volume 10.

[21] Smolin, L. (2009). *Generic predictions of quantum theories of gravity*, in *Approaches to quantum gravity*, ed. D.Oriti, Cambridge University Press, pp548-570.

[22] Sternberg, S. (1994). *Group Theory and Physics*. Cambridge University Press.

[23] Stillwell, J. (1993). *Classical Topology and Combinatorial Group Theory*, 2nd Edition, Springer Verlag.

[24] Thiemann, T. (2008). *Modern Canonical Quantum General Relativity*, Cambridge University Press.

[25] Wüthrich, C. (2017). *Raiders of the lost spacetime*, in *Towards a Theory of Spacetime Theories*, eds. D.Lehmkuhl, G.Schiemann, and E.Scholz, Birkhauser, pp 297-335.