FAQ on Canonical Loop Quantum Gravity

Generic questions

What is LQG?

Loop Quantum Gravity (LQG) is a proposed physical theory which describes quantum spacetime and the properties of gravitational interaction in the deep quantum regime. Its main features are:

- LQG is quantum: it admits precise (in the mathematical sense) formulation with a well-defined separable Hilbert space and self-adjoint operators corresponding to classical observables acting on the Hilbert space.
- LQG is a theory of gravity: according to numerical computations, it reproduces General Relativity in the classical limit.
- LQG is background independent: the formulation of the theory doesn't make any references to any pre-existing background spacetime structure.
- LQG is nonperturbative: it doesn't make use of assymptotic expansions in powers of the Newton's gravitational coupling constanit G.

What does it say about the world?

Among the unique predictions of Loop Quantum Gravity are the following:

- Spacetime is fundamentally discrete. For example, every physical measurement of the area of any surface (e.g. cross sections of interactions) gives a discrete value. The minimal possible value that the area of any surface could take is called the *area gap* Δ and is of order $\Delta \approx l_P^2$ where l_P is Planck's length.
- Singularities are artifacts of classical General Relativity and don't exist in the quantum theory. This is true of both black hole singularities and of the cosmological Big Bang singularity.
- Black hole entropy is proportional to its area, in agreement to the Bekenstein-Hawking formula. LQG provides an explanation for the holographic principle in terms of spin network links intersecting the holographic "screen".

What is background independence?

In the LQG circles it is widely believed that General Relativity is not just another field theory with a peculiar symmetry group. It is conceptually novel, that is, based on the *new fundamental principle* of background independence. Doing physics in the background independent setting is quite different from what we've been doing so far.

In simple words, background dependent theories like Maxwell's electrodynamics describe a form of substance (the electromagnetic field) living on the Minkowski spacetime of Special Relativity. These can be generalized to curved spacetimes of black holes or cosmology, but in any case the background structure is fixed. On the contrary, background independent theories like Einstein's General Relativity describe a substance (the gravitational field) which *makes up* spacetime. They don't make any references to any pre-existing structure, instead the possible spacetimes are given by solutions of the equations of motion.

Einstein's big discovery lies in the fact that spacetime and gravity is the same thing. It is often said in General Relativity that there is no gravity, but it is the curvature of spacetime that we observe as gravity; but it is also correct to say that there is no spacetime, but it is the properties of gravity that we observe as space and time. LQG takes this lesson of General Relativity seriously. The quantization of gravity in LQG is manifestly background independent.

What's wrong with perturbation theory?

Approximate methods are great tools to derive predictions. In almost all relevant cases the complexity of the system makes it is virtually impossible to do exact calculations. But there are complelling arguments against perturbation theory when it comes to gravity.

Perturbative approach to quantum gravity starts by expanding the metric around the flat Minkowski vacuum like so:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x).$$

Here $h_{\mu\nu}(x)$ can be treated as a basic field of the Quantum Field Theory of gravity living on the Minkowski background defined by $\eta_{\mu\nu}$. Let's put aside the aestetical argument against this split (it breaks manifest background independence) aside for a moment and consider the technical implications of using perturbation theory.

The first problem with this approach lies in the fact that the causal structure of the Quantum Field Theory of the field $h_{\mu\nu}(x)$ is given by the Minkowski background spacetime. For example, like in any QFT, the commutator of fields taken at spacelike separated points, has to be equal to zero (because otherwise it would be possible to transfer information through field correlations with superluminal speed). But the background metric $\eta_{\mu\nu}$ has no physical significance to the causal structure! The causal structure of General Relativity is determined by the full metric $g_{\mu\nu}(x)$, i.e. it is influenced by the field $h_{\mu\nu}(x)$, which is quantum. Thus we have an apparent conceptual contradiction between General Relativity and Quantum Field Theory already.

If we decide to ignore the first problem and carry on, we arrive at the second problem with perturbative Quantum Gravity: it is *nonrenormalizable*. This can be seen with a naked eye from simple dimensional analysis: the gravitational coupling constant (the Newton's constant $G = l_P^2/8\pi$) has negative dimensions of -2 in natural units, rendering the interaction *irrelevant* in the infrared (according to Wilson's classification) and uncontrollably divergent in the ultraviolet. This, of course, is a complete disaster for the perturbative quantization programme.

That nonrenormalizable theories can't be consistently quantized is a common misconception. Nonrenormalizability is unacceptable for any *perturbative* QFT, but there are famous examples of nonrenormalizable models being consistently quantized nonperturbatively. Probably the most famous (and relevant to us) example is General Relativity in 3 spacetime dimensions. The gravitational coupling in 3d has negative dimensions of -1 in natural units, and the perturbation theory is nonrenormalizable. However, 3d General Relativity has been consistently quantized nonperturbatively by Witten.

In LQG circles it is believed that nonrenormalizability is not the fault of General Relativity Lagrangian (like the case with Fermi's theory of weak interactions), but of the perturbative approach, which is inadequate when it comes to Quantum Gravity. That it is possible is proven a-posteriori, of course, by the existence of Loop Quantum Gravity. But even now it can be seen as follows: the uncontrollable growth of the gravitational coupling in the ultraviolet regime leads to the failure of the original assumption that the metric can be approximated by a perturbation around the flat Minkowski spacetime. At extremely short distances gravity dominated by violent quantum fluctuations of the structure of spacetime (aka Wheeler's "spacetime foam"), not by a smooth Minkowski solution. This is exactly the regime in which the high-frequency modes which give rise to ultraviolet divergences live. It is reasonable to assume that yet unknown short-scale nonperturbative effects of Quantum Gravity can effectively cut off the high-frequency modes and thus eliminate the divergences. This is exactly what happens in LQG, the mentioned effects having to do with fundamental *discretness* of spacetime which becomes important close to the Planck scale.

Is it Lorentz-invariant?

This question has been asked over and over for several decades. The source of the confusion lies in the existence of two inequivalent notions of Lorentz-invariance. As we will see, LQG is not different from General Relativity (GR) in both cases.

1. *Local Lorentz invariance* is satisfied when the theory is invariant under $SO(3,1)$ gauge transformations. For example, GR is invariant under local $SO(3,1)$ when formulated in terms of the frame field (aka tetrad or vierbein) $e^I_\mu(x)$. In this case the $SO(3,1)$ gauge transformations act on the internal (capital latin) indices.

2. *Global Lorentz invariance* is satisfied by the classical theory when there is a representation of $SO(3,1)$ which takes solutions of equations of motion to another solutions. For example, Maxwell's classical theory of electromagnetism and Dirac's theory of bispinor field are both globally Lorentz-invariant. Bosonic/supersymmetric strings in Minkowski target space are globally Lorentz invariant.
   As it turns out, classical General Relativity is not globally Lorentz-invariant. Those who disagree are welcome to describe the action of the Lorentz/Poincare groups on the cosmological FLRW solution of Einstein's equations in the comments. It is possible to select *a subset* of solutions with flat boundary on which the Lorentz transformations act globally. I am not sure, however, that these are relevant physically, since the cosmological solutions don't belong to this category.

Loop Quantum Gravity, just like classical General Relativity, is locally Lorentz-invariant, but not globally Lorentz-invariant. Global Lorentz invariance is just a property of the particular solution of Einstein's equations (Minkowski vacuum of Special Relativity).

But doesn't the existence of the area gap go against Lorentz invariance?

No, it doesn't. This is probably the most common misconception about LQG especially popular among superstring theorists.

The argument is usually formulated as follows: suppose we have an area gap (the minimal quantum of geometric area). If we pass to the moving reference frame, this area has to Lorentz-contract, and because by original assumption we don't have anything smaller than this area, we have a contradiction.

This argument is absolutely wrong. It fails because LQG is a quantum theory of spacetime, and the geometric properties of spacetime (like area or length) in it are also quantum. Discreteness of the spectrum of a quantum observable is known to coexist with continuous symmetries influencing the observable.

In simple words, consider a quantum state of spacetime geometry in which a surface has the minimal area Δ (the area gap). For convenience, we denote this state as $|\Delta\rangle$. What happens if we continuously Lorentz-contract it? The state changes, becoming a superposition of states with different areas of the surface:

$$|\Psi\rangle = \alpha\,|\Delta\rangle + \beta\,|0\rangle\,.$$

The mean area continuously decreases and becomes less than Δ:

$$\mathcal{A}(\Psi) = |\alpha|^2\,\Delta < \Delta.$$

There's no contradiction with the original assumption here, because it is the *eigenvalues* of area which are discrete.

What happens if we decide to measure the area of the surface? Quantum Mechanics gives an exhausted answer to this question: with probability $|\alpha|^2$ we will get the area gap Δ back, and with probability $|\beta|^2$ we will get zero (the surface disappears upon measurement).

Is it unitary?

Unitarity is required for the consistency of any Quantum Field Theory (QFT). LQG is not a QFT, and unitarity is not required for its consistency.

In QFT, time translations are symmetries of the theory (because they belong to the Poincare group). Any continuous symmetry has to be realized by a unitary operator on the Hilbert space of the quantum theory, because otherwise we could perform an experiment yielding statistical results, based on which we could distinguish between the initial and transformed systems, which we can't by the very definition of symmetry transformations.

In LQG, time translations *aren't* symmetries of the theory. In fact, they can't even be defined, because in LQG there is no external time in which the system evolves. This is a direct consequence of background independence. Time is a property of the gravitational field, and among all the other properties it has to be quantized, leaving no room for translations in external time.

Therefore, LQG doesn't even have operators corresponding to translations in external time, because it describes time on the equal footing with all the other properties of the gravitational field: its properties are encoded in the quantum state. Needless to say, it is meaningless to ask whether these are unitary or not.

More details on the nature of time in background independent quantum gravity and LQG in particular can be found later in this FAQ, under the section "Problem of time".

Does it unify interactions?

No.

LQG is a theory of Quantum Gravity. It can be coupled to different matter content, but it doesn't provide a machanism of unification. It was not intended as the final word in physics, but as another step. Its main achievement is the manifestly background independent quantum framework, which could in principle be used as a building block of the upcoming theory of everything :)

Is it elegant?

Elegance is a subjective concept, but many consider LQG extremely elegant.

Canonical General Relativity

Is LQG a quantization of some bare action?

It's called the Holst action for 4d General Relativity:

$$S_{\text{Holst}}[e,\omega] = \frac{1}{16\pi G}\int_M \underline{e}^I \wedge \underline{e}^J \wedge \left(* + \frac{1}{\gamma}\right) \underline{\underline{F}}_{IJ}(\omega) =$$

$$= \frac{1}{16\pi G}\int_M d^4x\ |e|\ e^\mu_I e^\nu_J\, P^{IJ}{}_{KL} F^{KL}_{\mu\nu}(\omega).$$

Here $*$ denotes the internal Hodge dual (and is metric-independent), $\underline{e}^I = e^I_\mu(x)dx^\mu$ is the frame form (aka tetrad or vierbein) and $\underline{\omega}^{IJ} = \omega^{IJ}_\mu(x)dx^\mu$ is the Lorentz connection. Note that $\underline{e}^I$ transforms trivially under $SO(3,1)$ gauge (local Lorentz) transformations:

$$\underline{e}^I \to g^I{}_J\, \underline{e}^J,$$

while the transformation law of the Lorentz connection is nonlinear:

$$\underline{\omega} \to g\underline{\omega}g^{-1} + gdg^{-1}.$$

Both fields transform as tensors under diffeomorphisms.

The curvature of the Lorentz connection is denoted as $2\underline{\underline{F}}^{IJ} = F^{IJ}_{\mu\nu}(x)\, dx^\mu \wedge dx^\nu$:

$$\underline{\underline{F}}^{IJ}(\omega) = d\underline{\omega}^{IJ} + \underline{\omega}^I{}_K \wedge \underline{\omega}^{KJ}.$$

The γ-deformed self-dual projector in the Holst action is defined as

$$P^{IJ}{}_{KL} = \delta^{[I}_K \delta^{J]}_L - \frac{1}{2\gamma}\epsilon^{IJ}{}_{KL}.$$

The *Immirzi* parameter γ is a dimensionless parameter which will play a major role in Loop Quantum Gravity. It doesn't influence the equations of motion, because the second term in the Holst action is *topological* (it is a full derivative and doesn't influence the equations of motion). Classicaly, the Holst action is equivalent to the Palatini action (obtained at $\gamma \to \infty$):

$$S_{\text{Palatini}}[e,\omega] = \frac{1}{64\pi G}\int_M \underline{e}^I \wedge \underline{e}^J \wedge \left(*\underline{\underline{F}}\right)_{IJ}(\omega) =$$

$$= \frac{1}{64\pi G}\int_M \epsilon_{IJKL}\, \underline{e}^I \wedge \underline{e}^J \wedge \underline{\underline{F}}^{KL}(\omega).$$

The equations of motion for both Palatini and Holst actions are thus:

1. The frame-connection compatibility condition:

$$D\underline{e}^I = d\underline{e}^I + \underline{\omega}^I{}_J \wedge \underline{e}^J = 0.$$

2. Einstein's equations in frame-connection formalism (for the imitation Ricci tensor):

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0,$$

$$R_{\mu\nu} = R^c{}_{acb} = F^{IJ}_{\mu\sigma} e^\sigma_I e^K_\nu \eta_{JK}.$$

What about Bianchi identities?

The imitation Bianchi identities take the form

$$D\underline{\underline{F}}^{IJ} = d\underline{\underline{F}}^{IJ} + \omega^I{}_K \wedge \underline{\underline{F}}^{KJ} = 0.$$

These follow from the properties of the covariant differential. First, we prove that $D^2 a = F \wedge a$:

$$D^2 a = D\left(da + \omega \wedge a\right) = d^2 a + d\left(\omega \wedge a\right) + \omega \wedge da + \omega \wedge \omega \wedge a =$$

$$= d\omega \wedge a - \omega \wedge da + \omega \wedge da + \omega \wedge \omega \wedge a = \left(d\omega + \omega \wedge \omega\right) \wedge a = F \wedge a.$$

Now consider the trivial identity:

$$D^3 a = D\left(D^2 a\right) = D^2\left(Da\right)$$

$$D\left(F \wedge a\right) = F \wedge Da$$

$$DF \wedge a + F \wedge Da = F \wedge Da$$

$$DF \wedge a = 0 \ \ (\forall a)$$

$$DF = 0.$$

Here we used the following property of differentials (valid for both d and D):

$$d(a \wedge b) = da \wedge b + (-1)^{\mathrm{deg} a} a \wedge db.$$

So, for example,

$$d\left(\omega \wedge a\right) = d\omega \wedge a - \omega \wedge da,$$

but

$$d\left(F \wedge a\right) = dF \wedge a + F \wedge da.$$

Why do we need spacetime foliation?

Spacetime foliation is the split of spacetime into hypersurfaces of equal coordinate time $t = \mathrm{const}$.

It is important to note that coordinate time t is unobservable parameter, which has nothing to do with the physical time. Contrary to some people feel, the split of spacetime into space and time doesn't go against manifest background independence.

In the modern spinfoam formulations the spatial hypersurfaces don't even have to be spacelike, corresponding to a choice of the *boundary* of the given spacetime region. The foliation thus reflects the nature of quantum mechanics: quantum states are associated to boundaries.

We use lower latin indices from the beginning of alphabet as manifold indices on the boundary: $a, b, \dots \in \{1, 2, 3\}$, and we use lower latin indices from the middle of the alphabet as internal $SO(3)$ indices: $i, j, \dots \in \{1, 2, 3\}$.

What is "The problem of time"?

General Relativity is an example of the *fully constrained system*. It means that its evolution is contained solely in constraints, while the Hamilton's function is trivially zero. Please see my derivation of constraints in General Relativity.

This is highly unusual because at the first sight it appears that the evolution is frozen implying the absense of change, which is contrary to our everyday-life experience. But actually, the problem of time is just another peculiar feature of background independent physics. Time is still there, but it is more complicated than a single axis.

The ADM Hamiltonian of General Relativity (in metric formulation) is a combination of four constraints:

$$H = \int_\Sigma d^3x \left(N(x) C + N_a(x) C^a\right),$$

where N and N^a are Lagrange multipliers called *lapse function* and *shift vector* respectively,

$$C^a = -2 D_b \pi^{ab}$$

are the 3d diffeomorphism constraints (the 3-metric q_{ab} on the boundary and its canonical momentum π^{ab} form the ADM phase space) and

$$C = \frac{1}{\sqrt{q}}\left(\left(q_{ac}q_{bd} - \frac{1}{2}q_{ab}q_{cd}\right)\pi^{ab}\pi^{cd} - \det q \cdot R(q)\right)$$

is the Hamiltonian constraint, which generates refoliations on shell.

So we start by a Cauchy data of

$$\{q_{ab}(\vec{x}),\, \pi^{ab}(\vec{x})\}$$

satisfying the 4 constraints:

$$C = C^a = 0.$$

How do we evolve it in time? What would be the 3d metric and its canonical momentum after 1 second passes?

This would be a valid question in any background dependent theory. But it is clearly silly in the background independent setting, where physical time itself depends on the gravitational field.

But we can choose a particular gauge by setting $N(\vec{x}) = 1$ and $N^a(\vec{x}) = 0$. Because we have (by ADM decomposition)

$$g_{00} = q^{ab}N_aN_b - N^2,$$

this choice of the lapse and shift corresponds to choosing coordinate time t to be equal to the physical time which has passed in the stationary frame of the foliated slice. We have thus a well-defined hamiltonian

$$H = \int_\Sigma d^3x \frac{1}{\sqrt{q}}\left(\left(q_{ac}q_{bd} - \frac{1}{2}q_{ab}q_{cd}\right)\pi^{ab}\pi^{cd} - \det q \cdot R(q)\right),$$

which vanishes on shell. But this doesn't mean that its Poisson brackets with q_{ab} and π^{ab} vanish! Thus we successfully recover the classical evolution from Hamilton's equations

$$\begin{cases} \dot{q}_{ab} = \{q_{ab},\, H\}, \\ \dot{\pi}^{ab} = \{\pi^{ab},\, H\}. \end{cases}$$

Note that we could've chosen another lapse and shift. The freedom of choice of N and N^a reflects the freedom to choose any coordinate system to describe the model. This is because the constraints which are enforced by the Lagrange multipliers generate diffeomorphisms and refoliations.

What are Ashtekar new variables?

After spacetime foliation, we partially fix the boost part of $SO(3,1)$. The relevant objects are

$$\Gamma_a^i = \frac{1}{2}\varepsilon^{ikl}\omega_a^{kl}, \quad K_a^i = \omega_a^{0i},$$

which completely determine the Lorentz connection (because $\underline{\omega}^{IJ}$ is antisymmetric in $I \leftrightarrow J$, the ω_a^{00} component is trivially zero).

Sometimes it is convenient to deal with

$$\Gamma_a^{ij} = \varepsilon^{ijk}\Gamma_a^k, \quad K_a^{ij} = \varepsilon^{ijk}K_a^k.$$

An important moment is the transformation properties of Γ_a^i and K_a^i under the rotation $SO(3)$ part of $SO(3,1)$. It follows from the structure of $SO(3,1)$ that Γ_a^{ij} transfoms as a connection under rotations $r^{ij}(x)$:

$$\Gamma_a \to r\Gamma_a r^{-1} + r\partial_a r^{-1},$$

whereas K_a^{ij} transforms as a vector:

$$K_a \to rK_a r^{-1}.$$

It is easy to see that if we add a vector to the $SO(3)$ connection, we get another $SO(3)$ connection. Consider the following connection:

$$A_a^i = \Gamma_a^i + \gamma K_a^i.$$

This connection arises as canonical conjugate to the densitized triad

$$\tilde{E}^a_i = \sqrt{\det q_{ab}} \cdot E^a_i = |E| \; E^a_i$$

in the Holst action:

$$\left\{\tilde{E}^a_i(x),\; A^j_b(y)\right\} = \gamma l_P^2 \delta^j_i \delta^a_b \cdot \delta^{(3)}(x,y);$$

$$\left\{\tilde{E}, \tilde{E}\right\} = \{A, A\} = 0.$$

where

$$l_P^2 = 8\pi G$$

is the Planck's area in natural units.

What is the significance of the Immirzi parameter?

As we already saw earlier, classical model described by the Holst action is independent of γ. The most logical choice is to set $\gamma \to \infty$, which gives the Palatini action for General Relativity in frame-connection variables. It also makes the secont term dominant in the Ashtekar-Barbero connection:

$$A^i_a = \Gamma^i_a + \gamma K^i_a \sim \gamma K^i_a.$$

The true conjugate of the inverse triad in the $\gamma \to \infty$ limit is thus K^i_a.

So why didn't we set $\gamma \to \infty$ from the start if this were to simplify our computations? Because the Immirzi parameter matters for the quantum theory. Upon quantization, different choices of γ lead to inequivalent theories, while the $\gamma \to \infty$ limit *does not exist*.

The crucial property of the Ashtekar-Barbero connection which allows us to quantize the theory is that *it transforms as an $SO(3)$ connection under gauge rotations*. This property will allow us to pass to the holonomy-flux algebra and quantize. It is impossible to use the LQG machinery on K^i_a because it is *not an $SO(3)$ connection*.

We will return to this later when we discuss the *holonomy-flux algebra*.

Loop quantization

What is holonomy?

Classical gauge geometry is given by the connection field $\underline{A}$ taking values in some Lie algebra $\mathfrak{g}$ of Lie group G. An important geometrical object associated to any path $\alpha : [0,1] \to \Sigma$ in the manifold Σ is the holonomy

$$h_A(\alpha) = \mathcal{P}\exp \int\limits_\alpha \underline{A} \in G.$$

We call a *generalized connection* a homomorphism from the group of open paths to G. Thus, a gauge connection $\underline{A}$ is also a generalized connection as it associates the holonomy $h_A(\alpha)$ for each path α. That it is a homomorphism is easy to check:

$$h_A(\alpha) h_A(\beta) = h_A(\alpha \circ \beta)$$

by the properties of path-ordered exponentials.

What is flux?

Flux is a geometric object naturally associated to the densitized triad $\tilde{E}^a_i$. For any surface $S \subset \Sigma$ the flux of $\tilde{E}^a_i$ through S is given by

$$\Phi_i = \int\limits_S d^2y \, \tilde{E}^a_i n_a,$$

where

$$n_a = \frac{1}{2}\epsilon_{abc}\epsilon^{\alpha\beta}\frac{\partial x^a}{\partial y^\alpha}\frac{\partial x^b}{\partial y^\beta}$$

is the normal to S in Σ.

Because the densitized triad is dual to a 2-form $2\underline{\underline{\Pi}}^i = \Pi^i_{ab}(x)\, dx^a \wedge dx^b$ by

$$\Pi^i_{ab} = \epsilon_{abc}\tilde{E}^a_i,$$

the integral for Φ_i can be written as an integral of the form over the surface S and is consequently coordinate-independent:

$$\Phi_i = \int_S \underline{\underline{\Pi}}^i.$$

How is flux related to area?

In background independent theories there is no a-priori notion of geometry. The surface $S \subset \Sigma$ doesn't have any "size" associated to it, in particular, its area is not defined. Instead, the area of the surface is an observable of the gravitational field.

In the previous section we've chosen a canonical pair $\{A^i_a,\, \tilde{E}^b_j\}$ (Ashtekar-Barbero connection and densitized triad) to span the phase space of General Relativity, because it is convenient for loop quantization. The question becomes: how can the area $\mathcal{A}(S)$ of the surface S be derived from the canonical pair?

The answer to this question is:

$$\mathcal{A}(S) = \int_S d^2y \sqrt{\delta^{ij} \cdot \tilde{E}^a_i n_a \cdot \tilde{E}^b_j n_b}.$$

It is easy to demonstrate: consider a local coordinate patch in which the first two coordinates lie in S and the third is normal to S. The metric in Σ takes the block-diagonal form

$$q_{ab} = \begin{pmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ 0 & 0 & q_{33} \end{pmatrix} = \begin{pmatrix} h_{\alpha\beta} & 0 \\ 0 & q_{33} \end{pmatrix}.$$

The inverse metric is

$$q^{ab} = \begin{pmatrix} h^{ab} & 0 \\ 0 & q^{33} \end{pmatrix}.$$

Consequently,

$$\det q^{-1} = \det h^{-1} \cdot q^{33},$$

or

$$\det h = \det q \cdot q^{33}.$$

The coordinate-independent expression for q^{33} is given by

$$q^{ab} n_a n_b$$

since it is equal to q^{33} in our coordinate system in which $n_a = \{0, 0, 1\}$ and is coordinate-independent. Thus, the area is given by

$$\mathcal{A}(S) = \int_S d^2y \sqrt{\det h} = \int_S d^2y \sqrt{\det q \cdot q^{ab} n_a n_b} = \int_S d^2y \sqrt{\delta^{ij} \cdot \tilde{E}^a_i n_a \cdot \tilde{E}^b_j n_b},$$

where we've used that for densitized triads

$$\delta^{ij} \tilde{E}^a_i \tilde{E}^b_j = \det q \cdot \delta^{ij} E^a_i E^b_j = \det q \cdot q^{ab}.$$

Another relation of extreme importance can be established between the area $\mathcal{A}(S)$ and the flux $\Phi_i(S)$. We split the surface S into infinitesimal chunks to give

$$\mathcal{A}(S) = \sum_x \sqrt{\delta^{ij} \cdot \Phi_i\left(\Delta S_x\right) \cdot \Phi_j\left(\Delta S_x\right)},$$

where x labels the chunk ΔS_x and $\Phi_i\left(\Delta S_x\right)$ is the flux of $\tilde{E}^a_i$ through the chunk.

What is and why do we need the holonomy-flux algebra?

Canonical quantization is based on implementing phase space coordinates as quantum-mechanical operators satisfying the canonical commutation relations. But canonical quantization doesn't do well when it comes to gravity. Actually, most of the Quantum Field Theories are also ill-defined when canonically quantized (see Haar's theorem for example).

Loop Quantum Gravity is based on a different quantization (so-called loop quantization), which makes use of another algebra: the *holonomy-flux algebra*. It can be derived formally by considering the poisson bracket between the Ashtekar-Barbero connection and the densitized triad. One needs the formula for the variation of the holonomy.

The resulting expression is

$$\{\Phi_i(S),\ h(\alpha)\} = \gamma l_P^2 \sum_x h(\alpha_1)\tau_i h(\alpha_2),$$

where the sum is over *oriented* intersections x of α and S, α_1 and α_2 being the holonomies of the two parts of the curve α separated by the intersection x.

What Hilbert space representation is used in LQG?

So far we have a conjugate pair of variables:

1. Ashtekar-Barbero connection $A_a^i(x)$ and its integral counterpart: holonomies of paths $h(\gamma)$;
2. Densitized triad $\tilde{E}_i^a$ and its integral counterpart: fluxes of surfaces $\Phi_i(S)$.

In the spirit of Quantum Mechanics we have to choose which one acts as multiplication when promoted to a self-adjoint operator on the Hilbert space, and which one acts as derivation. LQG is pretty specific about this:

Quantum states of gravity are functions of the Ashtekar-Barbero connection, which acts as multiplication. The conjugate densitized triad acts as derivation.

Formally we can represent states by functionals

$$|\Psi\rangle = \Psi[A].$$

The action of the Ashtekar-Barbero connection on the state is multiplicative:

$$\hat{A}_a^i(x)\,|\Psi\rangle = A_a^i(x)\,\Psi[A].$$

The action of the densitized triad is through derivation:

$$\hat{\tilde{E}}_i^a(x)\,|\Psi\rangle = -i\hbar\frac{\delta\Psi[A]}{\delta A_a^i(x)}.$$

This construction is, of course, rather formal. For example, we don't have a well-defined inner product. Formally we can write

$$\langle\Phi|\Psi\rangle = \int \mathcal{D}A\Phi^*[A]\Psi[A],$$

but no precise meaning to the functional measure $\mathcal{D}A$ can be given.

Is $\mathcal{D}A$ given by the Ashtekar-Lewandowski measure?

Yes!

Remember how the crucial fact which allows us to consistently quantize was that the Ashtekar-Barbero connection transforms as an $SU(2)$ connection? It was the reason to introduce the Holst term in the first place.

Well, this is why. There's a beautiful mathematical concept which gives a measure on so-called *cylindrical functions*, which are basically functionals of *generalized connections*.

LQG doesn't use Canonical Commutation Relations, nor does it use the fields A_a^i and $\tilde{E}_a^i$. In fact, it is crucial that these fields *can't* be represented as operators on the Hilbert space. Eventually this will lead us to the discreteness of spacetime at short scale.

Instead, LQG uses the holonomy-flux algebra of integral variables. Consider a graph of points (usually called nodes) joined by paths (usually called links). A cylindric function is the complex-valued function

$$\Psi(h_1, \ldots, h_n)$$

of the n holonomies associated to the links of the graph. The Ashtekar-Lewandowski measure on the space of cylindric functions is given by

$$\int \mathcal{D}A = \prod_{x=1}^{n} \int dh_x,$$

where we use the invariant Haar measure on $SU(2)$ for each link.

Isn't the Ashtekar-Lewandowski measure graph dependent?

At first sight a nasty artifact of graph dependence appears to plague this construction. However, it is not so: *the Ashtekar-Lewandowski measure is graph-independent.*

This can be shown as follows: consider a graph Γ and its subgraph Γ' (a subgraph consists of a subset of nodes and a subset of links). Each cylindrical function on Γ' can be represented as a cylindrical function on Γ which doesn't depend on the remaining links:

$$\Psi_{\Gamma'}(h_1, \ldots, h_{n'}) = \Psi_{\Gamma'}(h_1, \ldots, h_{n'}, \ldots, h_n).$$

The idea is then to choose a rich-enough graph from the start, knowing in advance that we won't need a richer structure later. And in case we do, just pretend that this structure was there from the start.

Since the Haar measure is normalized by

$$\int dh_x = 1,$$

this defines the graph independent Ashtekar-Levandowski measure on the space of cylindric functions.

Why are there different Hilbert spaces in LQG?

Well, LQG operates on different Hilbert spaces :)

The reason behind this diversity lies in the constraints. One has to correctly implement all the constraints of General Relativity as operators on the Hilbert space. The resulting Hilbert space of physical states is the mutual kernel of constraints (since physical states get annihilated by constraints).

The following spaces are used throughout LQG:

1. The space of cylindrical functions $\mathcal{H}_{\text{cyl}} = (\text{Cyl}, \mathcal{D}A)$ uses the Ashtekar-Lewandowski measure to define its inner product:

   $$\langle \Phi | \Psi \rangle = \prod_{x=1}^{n} \int dh_x \Phi^*(h_1, \ldots, h_n) \Psi(h_1, \ldots, h_n).$$

   An argument has been raised that $\mathcal{H}_{\text{cyl}}$ is nonseparable, which is true. However, the space of physical states, as long as the kinematical space of Loop Quantum Gravity — are both separable. Nonseparability is cured by background independence!

2. The space of gauge-invariant cylindrical functions $\mathcal{H}_{\text{gauge}}$ is obtained by requiring the functions to be $SU(2)$-invariant. Alternatively we could solve the Gauss constraint. The basis in $\mathcal{H}_{\text{gauge}}$ is given by *spin networks.*

3. The kinematical space $\mathcal{K}$ of LQG is the (dual of the) diffeomorphism-invariant $\mathcal{H}_{\text{diffeo}}$. The basis in $\mathcal{H}_{\text{diffeo}}$ is given by abstract *spin networks aka s-knots.* This space is separable, proving that all the excessive size of the original space was just gauge.

4. The physical space $\mathcal{H}$ is given by the kernel of the Hamiltonian constraint in $\mathcal{K}$. This is the only space to which we haven't been able to construct a convenient basis.

How exactly is spacetime discrete?

It is important to keep in mind that discreteness is not due to the preliminary setup, but due to quantization. It is fundamentally quantum.

It is often argued that the discreteness was put in the theory by considering cylindrical functions, but as we've seen it is not so: the Ashtekar-Lewandowski measure is graph independent. It allows all kinds of refinements to be made to the graph. Graphs are simply auxililary structures used to define the measure on the space of generalized connections.

How do we know that spacetime is discrete then? Well we could look at the spectrum of geometrical operators. Probably the simplest is the area operator. Classicaly, area can be determined by evaluating fluxes of the densitized triad through small chunks of the surface:

$$\mathcal{A}(S) = \sum_x \sqrt{\delta^{ij} \cdot \Phi_i\left(\Delta S_x\right) \cdot \Phi_j\left(\Delta S_x\right)}.$$

Can we promote this to a quantum operator acting on the kinematical Hilbert space $\mathcal{K}$? Yes!

First, let's describe how the area operator $\hat{\mathcal{A}}(S)$ acts on spin networks which intersect the surface S exactly once. Since the densitized triad acts through differentiation, we arrive at the direct consequence of the holonomy-flux algebra: the flux operator $\hat{\Phi}_i\left(\Delta S\right)$ inserts into the spin network state a single node located in the intersection point with

$$\gamma \hbar l_P^2 \cdot \tau_i$$

(where τ_i is the generator of $\mathfrak{su}(2)$) in the representation given by the link color. Since τ_i is not an intertwiner, the resulting state is not gauge-invariant, and thus doesn't live in $\mathcal{K}$, but rather in $\mathcal{H}_{\mathrm{cyl}}$.

However, the infinitesimal area operator $\delta\hat{\mathcal{A}} = \sqrt{\delta^{ij} \hat{\Phi}_i\left(\Delta S\right) \hat{\Phi}_j\left(\Delta S\right)}$ spawns two $\mathfrak{su}(2)$ generators τ_i and τ_j, which when contracted with the delta symbol give the quadratic Casimir of $\mathfrak{su}(2)$:

$$\delta^{ij} \tau_i \tau_j = j(j+1) \cdot \mathbb{I}.$$

This is proportional to the identity operator, which *is* a 2-valent intertwiner, and the resulting spin network is actually gauge-invariant and belongs to $\mathcal{K}$. Moreover, because the 2-valent intertwiner is proportional to the delta symbol, the resulting spin network is equivalent to the same original spin network. The only difference is in the numerical coefficient:

$$\delta\hat{\mathcal{A}} \left|\Psi\right\rangle = \gamma \hbar l_P^2 \sqrt{j(j+1)}.$$

In the general case, because area is additive, we can simply add contributions from all intersections with the spin network links:

$$\hat{\mathcal{A}}(S) = \gamma \hbar l_P^2 \cdot \sum_x \sqrt{j_x\left(j_x + 1\right)}.$$

This is the spectrum of area. For example, the smallest nonzero area that any surface can have (the area gap) is given by a single intersection with the spin-1/2 link:

$$\Delta = \frac{\sqrt{3}}{2} \gamma \hbar \cdot l_P^2.$$

Note that can't be fixed with refinement of the graph, because the discrete structure of the graph *is not* the cause of the spacetime discreteness. The real discreteness is in the spectra of geometrical observables.

Solving Hamiltonian constraint and spinfoams

There exist three approaches to solving the Hamiltonian constraint in LQG:

1. The canonical approach uses the classical redefinition of the Hamiltonian constraint

$$H = \int_\Sigma d^3x\, N \mathrm{tr}\left(F \wedge \{V,\, A\}\right),$$

   where $\hat{V}$ is the volume operator. This can be defined to act on spin networks. The kernel of this operator corresponds to the physical space of states $\mathcal{H}$.

2. The master constraint approach uses the single constraint

$$M = \int\limits_{\Sigma} d^3x \; (\mathrm{tr}\,(F \wedge \{V,\, A\}))^2$$

instead of the family of constraints parametrized by Lagrange multiplier $N(\vec{x})$.

3. The spinfoam approach will be the subject of the upcoming FAQ on spinfoams. It uses combinatoric generalizations of graphs (2-complexes) in order to model the *projection operator* on the kernel of the Hamiltonian constraint. It is still unknown if transition amplitudes in canonical and spinfoam pictures are the same, because no transition amplitudes have been obtained in the canonical picture yet. Spinfoam computations are much less opaque and much more results have been obtained.