AEI-2004-129
hep-th/0501114
January 14th, 2005

Loop quantum gravity: an outside view

Hermann Nicolai, Kasper Peeters and Marija Zamaklar

Max-Planck-Institut für Gravitationsphysik
Albert-Einstein-Institut
Am Mühlenberg 1
14476 Golm, GERMANY

hermann.nicolai, kasper.peeters, marija.zamaklar@aei.mpg.de

Abstract:

We present a pedagogical review of loop quantum gravity, with the aim of enabling a precise but critical assessment of its achievements so far. Special attention is paid to the appearance of a large number of ambiguities in the theory, in particular in the formulation of the Hamiltonian constraint. We emphasise that the off-shell ('strong') closure of the constraint algebra is a crucial test of the consistency of the theory, and should be used as the main tool to select one (if any) of the proposed Hamiltonians. Developing suitable approximation methods to establish a connection with classical gravity on the one hand, and with the physics of elementary particles on the other, remains a major challenge.

Contents

1 Key questions

When four-dimensional Einstein gravity is quantised canonically, in a perturbation series in the Newton constant around flat space-time, divergences arise at two-loop order. An impressive calculation by Goroff and Sagnotti [1, 2] has demonstrated that in order to obtain a finite S-matrix, the action should contain a counterterm

$$\Gamma^{(2)}_{\rm div} = \frac{1}{\epsilon}\frac{209}{2880}\frac{1}{(16\pi^2)^2}\int {\rm d}^4x\,\sqrt{g}\,C_{\mu\nu\rho\sigma}C^{\rho\sigma\lambda\tau}C_{\lambda\tau}{}^{\mu\nu}\,, \tag{1.1}$$

a result which was later confirmed in an independent background-field calculation by van de Ven [3]. The usual conclusion drawn from this result, and from the fact that the coupling constant is dimensionful, is that an infinite number of counterterms is needed. This non-renormalisability renders perturbatively quantised Einstein gravity useless as a fundamental theory since an infinite number of parameters would be required to make any physical prediction.

However, this still leaves open the possibility that Einstein gravity *can* be quantised consistently, but that it is simply the perturbation series in Newton's constant which is ill-defined.

This possibility has been raised and advocated not only in the context of canonical quantisation, but also, and independently, in the context of suggestions that there may exist a non-trivial fixed point of the renormalisation group in Einstein's theory [4, 5, 6]. Implicitly, it also underlies the path integral approach to Euclidean quantum gravity [7, 8], which provides a possible framework for the discussion of semi-classical states [9]. Indeed, to date there is no proof that such a quantisation which does *not* make use of a series expansion around a fixed background is guaranteed to fail. However, quantising without relying on perturbation theory around a free theory is hard.

Loop quantum gravity, or LQG for short, is an attempt to quantise Einstein gravity non-perturbatively. In contrast to string theory, which posits that the Einstein-Hilbert action is only an effective low energy approximation to some other, more fundamental, underlying theory, LQG takes Einstein's theory in *four* spacetime dimensions as the basic starting point. It makes use of many new technical ingredients. One of these are the Ashtekar variables, which (at least in the form originally proposed) greatly simplify the constraints. Another ingredient is the use of Wilson loops, or rather generalisations thereof, called spin networks. These in turn require other mathematical ingredients, such as non-separable ('polymer') Hilbert spaces and representations of operators, which fail to be weakly continuous (for explanation of these concepts see section 3). Although these new ingredients are mathematically interesting and perhaps necessary in order to circumvent the problems of perturbatively quantised gravity, it is important not to loose track of the physical questions that one is trying to answer.

The goal of the present paper is to review the essential properties of loop quantum gravity in an easily accessible way from a non-specialist's perspective, and with a non-LQG audience in mind. As there are already a number of excellent reviews available describing the story from the specialist's point of view [10, 11, 12, 13, 14], we will take the liberty to omit those mathematical details that in our opinion are not truly essential to understand the physical consequences of the formalism. Rather, after a summary of the 'prehistory' in section 2, we will try to describe and motivate some of the recent developments from an outside point of view. Accordingly, we will also take the liberty to explain some results 'our own way'. At the same time, as we move along, we will try to make precise and clearly state the questions that are often raised about the LQG programme (for earlier reviews which mention some of these concerns, see [15, 16, 17]).

In order to focus the discussion, and for the reader's convenience, we begin with a summary of what we consider to be the main open questions.

- *Structure of space(-time) at the smallest scales?*

  There is a general expectation (not only in the LQG community) that at the very shortest distances, the smooth geometry of Einstein's theory will be replaced by some quantum space or spacetime, and hence the continuum will be replaced by some 'discretuum'. LQG proposes a discrete structure of space [1], which is of a very peculiar type. More precisely, it is not the discreteness of a lattice or naive discretisation of space (*i.e.* of a finite or countable set). Rather, the 'discreteness' is imposed on the continuum by 'pulverising' it, *i.e.* by endowing it with the discrete topology, with countable unions of points as the open sets. Because the only notion of 'closeness' between two points in this topology is whether or not they are coincident, whence *any* function is continuous in this topology, this raises the question as to how one can recover conventional notions of continuity in this scheme. More specifically, is there any way of recovering familiar and well-established

[1] It is not immediately clear whether this would entail a discrete structure for *time* also, hence space-*time*. The reason is that there is no *a priori* notion of 'time' in quantum gravity. Instead, 'time' must be defined operationally in terms of a 'clock field', see e.g. [18, 19, 20]. Continuity or discreteness of time would then follow from the properties of that clock field and its spectrum.

concepts like the Wilsonian renormalisation group, where masses, couplings and correlators 'flow' *continuously* as some cutoff parameter is varied? The setup furthermore requires the *a priori* exclusion of certain 'ill-behaved' sets (such as infinite spin networks, that might contain Cantor or fractal sets), whose inclusion would lead to a breakdown of the formalism. These issues will be discussed in section 3.

- *How does smooth space-time appear in the classical limit?*

  The space of quantum states used in LQG (not necessarily a Hilbert space) is very different from the one used in Fock space quantisation. As a consequence, it becomes non-trivial to see how semi-classical 'coherent' states can be constructed, and how a smooth classical spacetime might emerge. In simple toy examples, such as the harmonic oscillator, it has been shown that the LQG quantisation method indeed leads to quantum states whose properties are close to those of the usual Fock space coherent states [21]. In full (3+1)-dimensional LQG, the classical limit is, however, far from understood (so far only kinematical coherent states are known [22, 23, 24, 25, 26, 27]). In particular, we do not know how to derive or even only describe the classical Einstein equations and their quantum corrections in this framework. A proper understanding of the semi-classical limit is also indispensable to understand the connection (or lack thereof) between conventional perturbation theory in terms of Feynman diagrams, and the non-perturbative quantisation proposed by LQG. For more on the structure of LQG Hilbert spaces, see section 3.1.

- *Renormalisation vs. regularisation: where is the 2-loop divergence?*

  The result (1.1) raises several questions as to what happens when one expands the results of non-perturbatively quantised Einstein gravity in Newton's constant. When such an expansion is performed about a semiclassical state (which remains to be found, see above), the two-loop divergence should manifest itself in one form or another. Or does there appear a cut-off (regulator) which eventually makes the perturbation theory finite? If so, where does this cut-off come from? Is it dynamically generated, like in string theory, or put in by hand? There are also questions concerning the meaning of 'regularisation'. According to conventional (quantum field theory) wisdom, physics is supposed not to depend on the way in which the theory is regulated before the cutoff is removed; how can it be that physics predictions of LQG do depend on the chosen regularisation prescription? This question is in part answered by the fact that the notions of 'finiteness' and 'regulator independence' as currently used in LQG on the one hand, and in conventional quantum field theory and perturbative quantum gravity on the other, are *not* the same; see section 3.5.

- *Background independence?*

  In a diffeomorphism invariant theory, there is no reference metric, hence no background measure of distance and scale, and this is generally thought to be a main distinctive feature of quantum gravity. LQG provides a background independent description of quantum gravity, which does not yet exist for string theory. However, it must be emphasised that 'background independence' here, at least so far, refers only to *spatial* backgrounds and *spatial* diffeomorphisms, and that, at the formal level, this property might also be claimed for the conventional geometrodynamics approach. The more challenging task is to make the formulation manifestly independent of a particular space-*time* background. This is notoriously difficult with any Hamiltonian approach. Let us furthermore stress that in spite of its perturbative origin, the result (1.1) cannot be so easily dismissed as a background artifact: while it does require *some* background for its derivation (*i.e.* a spacetime solving

Einstein's equations), the counterterm is actually independent of the particular background about which one expands, see [3], as is also evident from the manifestly space-time covariant form in which it is written. Concerning the completely different ways in which space and time are treated in LQG, compare section 4.1 with section 4.2.

- *Status of the Hamiltonian constraint?*

  In the current LQG literature, there is surprisingly little discussion of certain basic aspects concerning the Hamiltonian constraint operator, which we believe to be of central importance for the theory (but see e.g. [28] for a more recent treatment). For this reason, we will here describe the Hamiltonian constraint operator and its action on a given spin network wave function in rather pedestrian detail, as far as we have been able to work it out (the more difficult task of working out this action *explicitly* on the proper 'habitat' of states, a distribution space, appears to be beyond reach). In particular, we will exhibit the numerous choices and ambiguities inherent in this construction, as well as the extraordinary complexity of the resulting expression for the constraint operator — which makes us wonder what has really been gained in comparison with the old geometrodynamics approach, despite considerable technical progress. With regard to the myriads of possibilities, LQG proponents often express the view that these correspond to *different* physics, and therefore the choice of the correct Hamiltonian is ultimately a matter of physics (experiment?), and not mathematics. We disagree, because we cannot believe that Nature will allow such a great degree of arbitrariness at its most fundamental level. Instead, we are convinced that for the programme to succeed, additional consistency criteria (such as the proper closure of the quantum constraint algebra, see below) are badly needed. See section 4.2.

- *Does the quantum theory possess full spacetime covariance?*

  Spacetime covariance is a central property of Einstein's theory. Although the Hamiltonian formulation is not manifestly covariant, full covariance is still present in the classical theory, albeit in a hidden form, via the classical (Poisson or Dirac) algebra of constraints. However, this is not necessarily so for the quantised theory. LQG treats the diffeomorphism constraint and the Hamiltonian constraint in a very different manner. Why and how then should one expect such a theory to recover full spacetime (as opposed to purely spatial) covariance? The crucial issue here is clearly what LQG has to say about the quantum algebra of constraints. Unfortunately, to the best of our knowledge, the calculation of the commutator of two Hamiltonian constraints in LQG (with an explicit operatorial expression as the final result) has never been fully carried out. Instead, a survey of the possible terms arising in this computation has led to the conclusion that the commutator vanishes on a sufficiently large 'habitat' of states [29, 30], and therefore the LQG constraint algebra closes without anomalies. By contrast, we will here argue that this 'on shell closure' is not sufficient, but that a proper theory of quantum gravity requires a constraint algebra that closes 'off shell'. The fallacies that may ensue if one does not insist on off-shell closure can be illustrated with simple examples. In our opinion, this requirement may well provide the acid test on which LQG, like any other proposed theory of canonical quantum gravity, will stand or fail. See section 5.

- *Matter couplings: anything goes?*

  Because LQG is claimed to be a finite and fully consistent theory of quantum gravity, it does not appear to impose any restrictions on the types of matter that are coupled to gravity, nor on their interactions. Indeed it is straightforward, though sometimes cumbersome,

to extend the formalism to include matter: in this perspective, matter appears to be a mere accessory that can be added on to pure gravity as one chooses. This is in marked contrast to supergravity and superstring theory, which are based on the hypothesis that the very *raison d'etre* of matter is its indispensability for curing the perturbative (and non-perturbative) inconsistencies of quantum gravity, and the desire to 'geometrise' matter in the framework of a (probably supersymmetric) unified theory. It is difficult to see how LQG can recover the consistency requirements that conventional perturbative quantum field theory imposes on the matter content of the world, in particular those resulting from cancellation of *local (gauge) anomalies*. Let us recall that Nature does 'care' about such consistency requirements, in that it has chosen to put the three known generations of fermions into *anomaly free multiplets* of the standard model gauge group. Similar comments apply to global anomalies. Is it possible to obtain the correct answer for pion decay when fermions are coupled to electromagnetism in the LQG approach, or would LQG predict the neutral π meson to be a stable particle? See section 3.6.

Let us conclude this introduction with a list of items that we cannot cover in this review for lack of space, or lack of expertise on our side. These also represent important issues in quantum gravity, but would not essentially alter our assessment.

- *Quantum Cosmology*

  Much attention has been devoted recently to symmetry reduced versions of LQG as models of quantum cosmology [31, 32]. Apart from the possible avoidance of the big bang singularity, an appealing feature of these models is the possibility that inflation might be triggered and eventually stopped (gracefully) by gravity itself, via an intrinsically quantum gravitational mechanism [33]. In this way, the inflaton potential engineering that is characteristic of most current models of inflation could become unnecessary. However, while inspired by LQG in many ways, it is by no means clear whether and how these models can be derived from full-fledged LQG. For instance, the spectrum of the volume operator is not known, nor is it known whether its inverse is bounded.[2] Furthermore, the regularisation ambiguities of LQG feed through to the symmetry reduced models (such that e.g. the value of the Barbero-Immirzi parameter gets linked to the duration of the inflationary period).

- *Microscopic origin of black hole entropy*

  The explanation of the Bekenstein-Hawking entropy of black holes [34, 35] in terms of microstates has been claimed as a success for both LQG [36] and string theory [37]. The main achievement of string theory is that it not only explains the area law, but also predicts the factor 1/4 relating entropy and area; however, the argument requires a huge extrapolation in the string coupling constant, and is essentially limited to BPS-type extremal black holes. The LQG explanation, on the other hand, works for ordinary (Schwarzschild and Kerr) black holes, but relies on some assumptions that might be questioned (such as the use of Boltzmann rather than Bose statistics for the microstates), and requires a 'fit' for the Barbero Immirzi parameter to get the prefactor right. Although the two ansätze thus both reproduce the desired result, they seem almost impossible to reconcile, given the very different hypotheses underlying them — pure gravity on the one hand, and an exponentially growing spectrum of D-brane states on the other hand.

[2] There are indications that the spectrum of the inverse volume operator in the full theory is in fact not bounded from above (Thomas Thiemann, private communication).

- *Spin foams*

  Attempts to overcome the difficulties with the Hamiltonian constraint have led to yet another development, *spin foam models* [38, 39, 40]. These can be regarded as space-time versions of spin networks, to wit, evolutions of spin networks in 'time'. Mathematically, these models represent a generalisation of spin networks, in the sense that group theoretical objects (holonomies, representations, intertwiners, etc.) are attached not only to vertices and edges (links), but also to higher dimensional faces in a simplicial decomposition of space-time. Interesting as they are, however, these developments have so far not shed much new light on the problems with the Hamiltonian constraint, because decisive proof of the connection between these models (usually in the form 'BF-type' topological theories supplemented by extra constraints) and the full Einstein theory and its canonical formulation still appears to be lacking.

One final word concerning referencing: we have tried to give credit where credit is due, but we apologise in advance to any authors whose work we may have overlooked or underrepresented.

2 Old vs. new variables: from metric to loops

2.1 Prelude: the metric (or dreibein) approach

There are many introductory texts to which we can refer readers for a more detailed treatment of canonical gravity (see e.g. [12, 16, 19, 41, 42, 43, 44, 45]), but let us nevertheless briefly review the traditional way of doing canonical quantum gravity, also called *geometrodynamics*. Our exposition in this section follows [46], whose notations and conventions we adopt in the remainder. We will be using the vierbein formalism with a vierbein ${E_\mu}^A$ and space-time metric $G_{\mu\nu} = {E_\mu}^A {E_\nu}^B \eta_{AB}$ with flat (tangent space) metric $\eta_{AB} = \mathrm{diag}\,(-1,+1,+1,+1)$ [3]. The vierbein is covariantly constant under a derivative which is covariant w.r.t. both spacetime diffeomorphisms and local Lorentz transformations, viz.

$$D_\mu {E_\nu}^A := \nabla_\mu {E_\nu}^A - {\omega_\mu}^{AB} E_{\nu B} = 0\,, \tag{2.1}$$

where ∇_μ is the covariant derivative which involves only the Christoffel connection,

$$\nabla_\mu V_\nu := \partial_\mu V_\nu - {\Gamma_{\mu\nu}}^\rho V_\rho \tag{2.2}$$

and $\omega_{\mu AB}$ is the spin connection. It is a standard result that both ${\Gamma_{\mu\nu}}^\rho$ and $\omega_{\mu AB}$ can be explicitly solved for the vielbein from the above equation (in the absence of torsion).

As is customary in canonical gravity, and following the standard ADM prescription [41] we assume the space-time manifold $\mathcal{M}$ to be foliated according to $\mathcal{M} = \Sigma \times R$, with a spatial manifold Σ of fixed topology (and no boundary, for simplicity). Using letters $m, n, ...$ and $a, b, ...$ for curved and flat *spatial* indices, respectively, we choose the triangular gauge for the vierbein

$${E_\mu}^A = \begin{pmatrix} N & N^a \\ 0 & {e_m}^a \end{pmatrix} \quad \Longrightarrow \quad G_{\mu\nu} = \begin{pmatrix} -N^2 + N_a N_a & N_n \\ N_m & g_{mn} \end{pmatrix}, \tag{2.3}$$

where $g_{mn} \equiv {e_m}^a e_{na}$ is the spatial metric. The vierbein components N and N_a are Lagrange multipliers ('lapse' and 'shift'); their variation yields the canonical constraints. In the remainder we will freely convert spatial indices by means of the spatial dreibein and its inverse.

[3] Modulo some dimension dependent factors, the results described in this subsection are valid in any dimension.

The canonical momenta are obtained in the standard fashion from the Einstein-Hilbert action

$$\Pi_a{}^m := \frac{\delta\mathcal{L}}{\delta\partial_t e_m{}^a} = \frac{1}{2}ee_b{}^m(K_{ab} - \delta_{ab}K)\,, \tag{2.4}$$

where $e \equiv \det(e_m{}^a)$, and

$$K_{ab} := E_a{}^\mu E_b{}^\nu \nabla_\mu E_{\nu 0} = \omega_{ab0} \tag{2.5}$$

is the extrinsic curvature of Σ (expressed in terms of flat spatial indices, with $\omega_{ab0} \equiv E_a{}^\mu \omega_{\mu b0}$), and $K \equiv K_{aa}$. The inverse formula reads

$$K_{ab} = \frac{2}{e}\big(\Pi_{ab} - \frac{1}{2}\delta_{ab}\Pi\big) \quad , \qquad (\Pi \equiv \Pi_{aa})\,. \tag{2.6}$$

From the symmetry $K_{ab} = K_{ba}$ we immediately deduce the Lorentz constraint

$$L_{ab} = e_{m[a}\Pi_{b]}{}^m \approx 0\,, \tag{2.7}$$

which is also the canonical generator of spatial rotations on the dreibein ($\approx$ means 'weakly zero' [43]). The canonical Hamiltonian (really Hamiltonian density) is

$$H = \partial_t e_a{}^m \Pi_m{}^a - \mathcal{L} = NH_0 + N_a H_a\,, \tag{2.8}$$

with the diffeomorphism constraint

$$H_a \equiv D_m \Pi_a{}^m \approx 0\,, \tag{2.9}$$

and the Hamiltonian constraint (*alias* the scalar constraint)

$$H_0 \equiv e^{-1}\Big(\Pi_{ab}\Pi_{ab} - \frac{1}{2}\Pi^2\Big) - eR^{(3)} = \frac{1}{4}e\big(K_{ab}K_{ab} - K^2\big) - eR^{(3)} \approx 0\,, \tag{2.10}$$

where $R^{(3)}$ is the spatial Ricci scalar. The canonical equal time (Poisson) brackets are

$$\{e_{ma}(\mathbf{x}), \Pi_b{}^n(\mathbf{y})\} = \delta_{ab}\delta_m^n\delta^{(3)}(\mathbf{x},\mathbf{y})\,, \tag{2.11}$$

with the other brackets vanishing (for the canonical variables with the indices in the indicated positions). Canonical quantisation in the 'position space representation' now proceeds by representing the dreibein as a multiplication operator, and the canonical momentum by the functional differential operator

$$\Pi_a{}^m(\mathbf{x}) = \frac{\hbar}{i}\frac{\delta}{\delta e_m{}^a(\mathbf{x})}\,. \tag{2.12}$$

With these replacements, the classical constraints are converted to quantum constraint operators which act on suitable wave functionals. The diffeomorphism and Lorentz constraints become

$$H_a(\mathbf{x})\Psi[e] = 0 \quad , \qquad L_{ab}(\mathbf{x})\Psi[e] = 0\,. \tag{2.13}$$

They will be referred to as 'kinematical constraints' throughout. Dynamics is generated via the Hamiltonian constraint, the *Wheeler-DeWitt (WDW) equation* [47, 48, 49, 50]:

$$H_0(\mathbf{x})\Psi[e] = 0\,. \tag{2.14}$$

It is straightforward to include matter degrees of freedom, in which case the constraint operators and the wave functional $\Psi[e, \dots]$ depend on further variables (indicated by dots). The functional $\Psi[e, \dots]$ is sometimes referred to as the 'wave function of the universe', and is supposed to contain the complete information about the universe 'from beginning to end'. A good way to visualise Ψ

is to think of it as a film reel; 'time' and the illusion that 'something happens' emerge only when the film is played.

The substitution (2.12) turns the Hamiltonian constraint into a highly singular functional differential equation, which most likely cannot be made mathematically well defined in this form, even allowing for certain 'renormalisations'. Here we do not wish to belabour the well known difficulties, both mathematical and conceptual, which have stymied progress with the WDW equation for over forty years, see e.g. [19, 42, 45] [4]. However, we would like to emphasise the (surely well known) fact that, at least formally and ignoring fine points of functional analysis (such as constructing scalar products and appropriate Hilbert spaces of wave functionals), solutions to both the Lorentz constraint and the diffeomorphism constraint can be rather easily obtained by integrating suitably densitised invariant combinations of the spatial dreibein (or metric) and curvature components, and the matter 'position variables' over the spatial three-manifold Σ. The associated wave functionals Ψ are then automatically functionals of *diffeomorphism classes* of spatial metrics. These simple observations already show that, beyond technical subtleties, the kinematical constraints are not the real problem of quantum gravity. The core difficulties of canonical quantum gravity are all connected in one way or another to the Hamiltonian constraint – irrespective of which canonical variables are used.

2.2 Ashtekar's new variables

Much of the initial excitement over Ashtekar's discovery [51] of new canonical variables was due to the change of perspective they bring about, which fuelled hopes that they might alleviate some of the longstanding unsolved problems of quantum gravity. Let us therefore first describe what they are, and how they are obtained. We will skip the systematic derivation here based on the addition of a term $\propto \int E \wedge E \wedge R$ to the Einstein-Hilbert action [52, 53] (this term vanishes upon use of the Bianchi identity; though trivially invariant under deformations of the metric, it is not associated with any topological charge). Instead consider the combination [5]

$$A_{ma} := -\frac{1}{2}\epsilon_{abc}\omega_{mbc} + \gamma e^{-1}(\Pi_{ma} - \frac{1}{2}e_{ma}\Pi) = -\frac{1}{2}\epsilon_{abc}\omega_{mbc} + \gamma K_{ma}\,, \tag{2.15}$$

where ω_{mbc} is the *spatial* spin connection. The parameter $\gamma \neq 0$ is referred to as the 'Barbero-Immirzi parameter' in the LQG literature [54, 55]. Classically, γ has no physical significance, but is believed to become physically relevant upon quantisation, for instance, by setting the scale for the fundamental areas and volumes (in this sense it is somewhat analogous [56] to the θ parameter of QCD). One can now show that

$$\begin{aligned} \{A_{ma}(\mathbf{x}),\ A_{nb}(\mathbf{y})\} &= 0\,, \\ \{\tilde{E}^m_a(\mathbf{x}),\ \tilde{E}^n_b(\mathbf{y})\} &= 0\,, \\ \{A_m{}^a(\mathbf{x})\,,\ \tilde{E}^n_b(\mathbf{y})\} &= \gamma\delta^n_m\delta^a_b\delta^{(3)}(\mathbf{x},\mathbf{y})\,, \end{aligned} \tag{2.16}$$

where the canonically conjugate variable to A_{ma} is the *inverse densitised spatial dreibein*

$$\tilde{E}^m_a := e e_a{}^m \qquad \Longrightarrow \quad \tilde{E} = e^2\,. \tag{2.17}$$

with $e \equiv \det e_m{}^a$ and $\tilde{E} \equiv \det \tilde{E}^m_a$. The parameter γ is often eliminated from these brackets by absorbing it into the definition of $\tilde{E}^m_a$.

[4] One of the inventors of (2.14) has been overheard aptly referring to it as *"that damned equation..."*.

[5] We alert readers that our notational conventions differ from the ones used in most of the LQG literature, where $a, b, \ldots$ denote curved and $i, j, \ldots$ flat spatial indices.

To rewrite the constraints in terms of the new variables, we first observe that the covariant constancy of the spatial dreibein and the Lorentz constraint imply the Gauss constraint:

$$D_m \tilde{E}^m_a \equiv \partial_m \tilde{E}^m_a + \epsilon_{abc} A_m{}^b \tilde{E}^{cm} \approx 0\,. \tag{2.18}$$

Defining the field strength

$$\begin{aligned} F_{mna} &:= \partial_m A_{na} - \partial_n A_{ma} + \epsilon_{abc} A_{mb} A_{nc} \\ &= -\frac{1}{2}\epsilon_{abc} R_{mnbc} + \gamma\big(D_m K_{na} - D_n K_{ma}\big) + \gamma^2 \epsilon_{abc} K_m{}^b K_n{}^c\,, \end{aligned} \tag{2.19}$$

it follows that the diffeomorphism constraint takes the form

$$\tilde{E}^m_a F_{mna} \approx 0\,. \tag{2.20}$$

Furthermore,

$$\begin{aligned} \epsilon_{abc} \tilde{E}^m_a \tilde{E}^n_b F_{mnc} &= -\gamma^2\big(\Pi_{ab}\Pi_{ab} - \frac{1}{2}\Pi^2\big) - e^2 R^{(3)} \\ &= -\gamma^2 e \mathcal{H}_0 - \frac{1}{4}(1+\gamma^2) e^2 (K_{ab}K_{ab} - K^2)\,. \end{aligned} \tag{2.21}$$

These relations immediately suggest interpreting $A_m{}^a$ as a *gauge connection* for the gauge group $SO(3)$ of spatial rotations (for $D = 2+1$ gravity, this group is replaced by its non-compact form $SO(1,2)$ [57]). Accordingly, the new variables are conveniently rewritten as [51]

$$A_{m\alpha\beta} \equiv A_{ma} \tau^a_{\alpha\beta}\,, \tag{2.22}$$

where τ^a are the Pauli matrices.

For the special choice $\gamma = \pm i$ [51] the second term on the r.h.s. of (2.21) drops out, and — save for an extra density factor of e — the Hamiltonian constraint is expressed entirely in terms of the new canonical variables, and furthermore depends on them *polynomially*. In other words, this particular choice allows us to combine the diffeomorphism and Hamiltonian constraints, which are schematically of the form '$(\partial + \omega)\Pi \approx 0$' and '$(\partial\omega + \omega^2 + \Pi^2) \approx 0$', respectively, into a single expression in terms of the connection $A = \omega + \Pi$, and its canonically conjugate variable. Moreover, for this choice of γ the connection $A^{(\pm)}_{ma}$ is nothing but the pullback of the four-dimensional spin connection to the spatial hypersurface Σ, with the two signs corresponding to the two chiralities (indicated by superscripts $^{(\pm)}$)

$$\omega_{mAB}\gamma^{AB} = \omega_{mab}\gamma^{ab} + 2\omega_{ma0}\gamma^a\gamma^0 = A^{(+)}_{ma}\gamma^0\gamma^a(1+\gamma^5) + A^{(-)}_{ma}\gamma^0\gamma^a(1-\gamma^5)\,, \tag{2.23}$$

(cf. eqn. (2.5)). Equivalently, the variables $A^{(\pm)}_{ma}$ are associated with the selfdual and anti-selfdual parts of the spin connection, respectively,

$$A^{(\pm)}_{ma} = \omega^{(\pm)}_{ma0} \qquad \text{with} \qquad \omega^{(\pm)}_{mAB} := \frac{1}{2}\left(\omega_{mAB} \pm \frac{i}{2}\epsilon_{ABCD}\omega_m{}^{CD}\right)\,. \tag{2.24}$$

This is therefore also the natural choice for γ when one considers coupling gravity to chiral fermions. In fact, one of the authors (H.N.) was first enticed into learning about Ashtekar's variables when realising that they simplify the calculation of the constraint algebra of supergravity considerably [58, 59]. This is because the local supersymmetry constraint (*i.e.* the time component of the Rarita-Schwinger equation) always contains a factor $D_{[m}(A)\psi_{n]}$, where ψ_n is the gravitino, and A is just the Ashtekar connection (of course with $\gamma = \pm i$). More succinctly, in supergravity, the commutation relations (2.16), as well as the polynomiality of the constraints must necessarily

hold, if the commutator of two local supersymmetry constraints is to close into the scalar and diffeomorphism (and possibly other) constraints.

From the esthetical point of view $\gamma = \pm i$ is therefore clearly the preferred choice. Nevertheless, this value has been abandoned in most of the recent LQG literature, because there is a major difficulty with it: the phase space of general relativity must be *complexified* with imaginary or complex γ. To recover the *real* phase space of general relativity and to ensure that real initial data evolve into real solutions, suitable reality conditions must be imposed. This is straightforward to achieve for the classical theory — after all, we have merely changed the variables, not the theory itself — but not so for the quantum theory. There, the complexification poses subtle problems concerning the definition and imposition of appropriate hermiticity conditions on the states and operators, and no consensus has been reached so far on how to circumvent these difficulties (or on whether they can be circumvented at all).

For this reason one now usually takes γ to be *real*. In this case, no problem arises with reality of solutions in either the classical or the quantum theory, but the polynomiality of the constraints, and hence one of the most attractive features of the new variables, is lost (but let us note that problems with reality may re-surface when fermionic couplings are included, even if they are absent in the bosonic sector [59]). This is because the extra term in (2.21) no longer vanishes, but must be subtracted from both sides to recover the correct WDW Hamiltonian. Accordingly, one must express the extra contributions in terms of the new canonical variables ${A_m}^a$ and $\tilde{E}^m_a$ via (2.15) and (2.17), and this in turn requires expressing the original dreibein as well as the extrinsic curvatures in terms of the new canonical variables. For a while this was regarded as a chief obstacle, until a way to solve it was discovered by Thiemann [60]. To this aim let us first introduce the volume associated with a region $\Omega \subset \Sigma$ (considered as a phase space variable) [6]

$$V(\Omega) = \int_\Omega \mathrm{d}^3x\, e = \int_\Omega \mathrm{d}^3x\, \sqrt{\tilde{E}} \equiv \int_\Omega \mathrm{d}^3x\, \sqrt{\frac{1}{3!}\epsilon_{mnp}\epsilon^{abc}\tilde{E}^m_a\tilde{E}^n_b\tilde{E}^p_c}\,. \tag{2.25}$$

Writing $V \equiv V(\Sigma)$, we first use the substitution

$${e_m}^a(\mathbf{x}) = \epsilon_{mnp}\epsilon^{abc}\tilde{E}^{-1/2}\tilde{E}^n_b\tilde{E}^p_c(\mathbf{x}) = \frac{1}{4\gamma}\Big\{{A_m}^a(\mathbf{x}), V\Big\} \tag{2.26}$$

to recover the spatial dreibein. The second trick is to eliminate the extrinsic curvature using a doubly nested bracket. The first bracket is introduced by rewriting

$${K_m}^a(\mathbf{x}) = \frac{1}{\gamma}\big\{{A_m}^a(\mathbf{x})\,,\,\bar{K}\big\} \qquad \text{where} \qquad \bar{K} \equiv \bar{K}(\Sigma) := \int_\Sigma \mathrm{d}^3x\, {K_m}^a\tilde{E}_a{}^m\,. \tag{2.27}$$

The second bracket comes in through identity

$$\bar{K}(\mathbf{x}) = \frac{1}{\gamma^{3/2}}\Big\{\frac{\tilde{E}_a{}^m\tilde{E}_b{}^n}{\sqrt{\tilde{E}}}\epsilon^{abc}F_{mnc}(\mathbf{x})\,,\,V\Big\}\,. \tag{2.28}$$

Last, we need another dreibein factor to convert the curved index m on ${K_m}^a$ to a flat one, for which we must use (2.26) once again. We will postpone writing out the Hamiltonian constraint in terms of these multiple brackets until section 4.2. But let us note already here that the above relations supply the 'tool kit' also for transcribing any given matter Hamiltonian in terms of the new variables.

For $\gamma \neq \pm i$ the connection is no longer the pullback of a four-dimensional spin connection [61]. Although this does not immediately lead to problems with spacetime covariance, we will see in section 3.6 that the problem comes back when one considers coupling the theory to fermionic matter degrees of freedom.

[6] The ϵ-symbol is always the invariant *tensor density*: $\epsilon_{mnp} = e^{-1}{e_m}^a{e_n}^b{e_p}^c\epsilon_{abc}$, *i.e.* assumes the values $0, \pm 1$.

2.3 The connection representation

Independently of the choice of γ, the reformulation of canonical gravity in terms of connection variables opens many new avenues, in particular the use of concepts, tools and techniques from Yang-Mills theory. Early attempts at quantisation were based on the *connection representation* with the original Ashtekar connection, *i.e.* $\gamma = \pm i$. Although these were ultimately not successful, let us nonetheless briefly summarise them here. In this scheme, one represents the connection ${A_m}^a$ by a multiplication operator, and sets

$$\tilde{E}^m_a(\mathbf{x}) = \frac{\hbar}{i}\frac{\delta}{\delta {A_m}^a(\mathbf{x})}\,. \tag{2.29}$$

The WDW functional depending on the spatial metric (or dreibein) is accordingly replaced by a functional $\Psi[A]$ living on the space of connections (modulo gauge transformations). The price one pays is that this representation is much harder to 'visualise' because the spatial metric is no longer represented by a simple multiplication operator, but must now be determined from the operator for the inverse densitised metric

$$g g^{mn}(\mathbf{x}) = -\hbar^2 \frac{\delta}{\delta {A_m}^a(\mathbf{x})}\frac{\delta}{\delta {A_n}^a(\mathbf{x})} \quad . \tag{2.30}$$

Even if one ignores the clash of functional differential operators at coincident points, finding suitable states and computing their expectation values is obviously not an easy task (and has not been accomplished so far). Similarly, the spatial volume density is obtained from

$$g(\mathbf{x}) = \tilde{E}(\mathbf{x}) = \frac{i\hbar^3}{6}\epsilon^{abc}\epsilon_{mnp}\frac{\delta}{\delta {A_m}^a(\mathbf{x})}\frac{\delta}{\delta {A_n}^b(\mathbf{x})}\frac{\delta}{\delta {A_p}^c(\mathbf{x})} \quad . \tag{2.31}$$

Again this operator is very singular. Equations (2.30) and (2.31) provide a first glimpse of the difficulties that the formalism has in finding semi-classical states and thereby establishing a link between the quantum theory and classical smooth spacetime geometry.

For the quantum constraints the replacement of the metric by connection variables leads to a Hamiltonian which is simpler than the original WDW Hamiltonian, but still very singular. Allowing for an extra factor of e (and assuming $e \neq 0, \infty$) the WDW equation becomes

$$\epsilon^{abc} F_{mna}\big(A(\mathbf{x})\big)\frac{\delta}{\delta {A_m}^b(\mathbf{x})}\frac{\delta}{\delta {A_n}^c(\mathbf{x})}\Psi[A] = 0\,. \tag{2.32}$$

Here we have adopted a particular ordering, which however is by no means singled out. No viable solutions to this constraint have been found, but there is at least one interesting solution if one allows for a non-vanishing cosmological constant Λ. Namely, using an ordering opposite to the one above, and including a term Λg with the volume density (2.31), the WDW equation reads

$$\epsilon_{abc}\frac{\delta}{\delta {A_m}^a(\mathbf{x})}\frac{\delta}{\delta {A_n}^b(\mathbf{x})}\left(F_{mnc}\big(A(\mathbf{x})\big) - \frac{i\hbar\Lambda}{6}\epsilon_{mnp}\frac{\delta}{\delta {A_p}^c(\mathbf{x})}\right)\Psi_\Lambda[A] = 0\,. \tag{2.33}$$

This is solved by

$$\Psi[A] = \exp\left(\frac{i}{\hbar\Lambda}\int_\Sigma \mathrm{d}^3\mathbf{x}\,\mathcal{L}_{CS}(A)\right)\,, \tag{2.34}$$

with the Chern-Simons Lagrangian $\mathcal{L}_{\mathrm{CS}} = A\wedge \mathrm{d}A + iA\wedge A\wedge A$. In the literature this state is known as the Kodama state [62], but the solution had been known for a long time in Yang-Mills theory [63] (where however it has rather unusual physical properties [64]). The difficulties with

this solution have been much discussed recently [64, 65]; an obvious one being that the flat space limit $\Lambda \to 0$ does not appear to exist.

What happens when we choose γ to be different from $\pm i$, and real in particular? As we explained already, the extra term in (2.21) then no longer vanishes, must be dealt with separately [60]. The nice polynomial form of the Hamiltonian constraint operator (2.32) is lost. When implementing the translation rules at the end of the foregoing subsection in the connection formulation, one finds that the new Hamiltonian is not significantly simpler any more than the original one of geometrodynamics in terms of metric or dreibein variables.

2.4 From connections to holonomies

The *loop representation* is an attempt to overcome the difficulties with the connection representation which we sketched above. The transition between the connection and the loop representation was originally obtained via the *loop transform*, which can be thought of as a kind of functional Fourier transform [66]. We will not describe that construction here, but turn immediately to the formulation in terms of holonomies, on which the modern formulation of LQG – spin networks and spin foams – are based.

Whereas in the connection representation one works with functionals $\Psi[A]$ which are supported 'on all of Σ', one now switches to the *holonomies* as the basic variables. These are gauge covariant functionals supported on one-dimensional links, or 'edges', which we will designate by e (following established LQG notation). For a given edge, *i.e.* some (open) curve embedded in Σ, we set

$$h_e[A] = \mathcal{P} \exp \int_e A_m \mathrm{d}x^m \,, \quad \text{with } A \equiv A^a \tau_a \,. \tag{2.35}$$

Hence, $h_e[A]$ is a matrix valued functional. The holonomy transforms under the action of SU(2) at each end of the edge e:

$$h_e[A] \to h_e^g[A] = g\big(e(0)\big)\; h_e[A]\; g^{-1}\big(e(1)\big) \,, \qquad \text{with} \quad g\big(e(0)\big) \,,\; g\big(e(1)\big) \;\in \mathrm{SU}(2) \,. \tag{2.36}$$

For the remainder it is important that the holonomies are to be regarded as variables *in their own right*, subject to these transformation properties (so in some sense one can 'forget' about the original connection A defining the holonomy). The *distributional nature* of the holonomy is not only evident from its singular support (on a line rather than all of Σ), but also from the fact that we do not assume $h_e[A]$ to be close to the identity if the edge e is 'small' (this terminology has to be used with due care, as there is no a priori measure that tells us when e is 'small', but we can still imagine making it 'small' by chopping e into as many 'subedges' as we like). The fact that the typical field configuration is generically a distribution rather than some smooth function is well known from constructive quantum field theory [67].

The holonomies are taken to transform in SU(2) representations ρ_{j_e} of *arbitrary spin* $j_e = \frac{1}{2}, 1, \frac{3}{2}, \ldots$ for each link e (with the convention that $j_e = 0$ means that there is no edge). We will denote such a spin-j_e valued holonomy by

$$\big(\rho_{j_e}(h_e[A])\big)_{\alpha\beta} \,, \tag{2.37}$$

with indices $\alpha, \beta, \ldots$ as appropriate for the representation at hand. To make the notation less cumbersome, we will occasionally suppress ρ_{j_e} and the representation labels, and simply denote the above matrix as $(h_e[A])_{\alpha\beta}$.

To define the conjugate variable, we recall that the area element for the spatial manifold Σ can be expressed as a Lorentz vector (*i.e.* with flat $SO(3)$ indices) via

$$\mathrm{d}F_a := \epsilon_{abc} \theta^b \wedge \theta^c \qquad \text{with} \quad \theta^a \equiv \mathrm{d}x^m e_m{}^a \,. \tag{2.38}$$

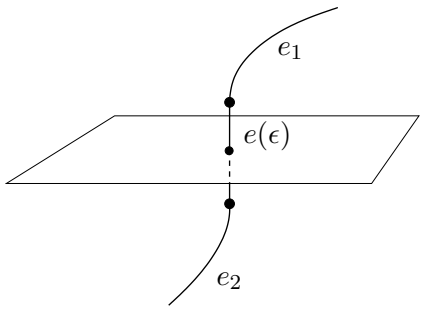

Figure 1: Setup used for the computation of the bracket (2.42). In the limit $\epsilon \to 0$ the edge $e(\epsilon)$ shrinks to zero and the two nodes just above and below the surface coincide.

Happily, this can be nicely rewritten in terms of the new canonical variables

$$\mathrm{d}F_a = \epsilon_{abc} e_n{}^b e_p{}^c \mathrm{d}x^n \wedge \mathrm{d}x^p = \epsilon_{mnp} \tilde{E}^m_a \mathrm{d}x^n \wedge \mathrm{d}x^p \,. \tag{2.39}$$

As the conjugate variable to $h_e[A]$ one takes the 'flux' vector

$$F^a_S[\tilde{E}] := \int_S \mathrm{d}F^a \tag{2.40}$$

through any two-dimensional surface S embedded in Σ. There is also a smeared version of this variable, with a test function f_a to soak up the free index a, which reads

$$F_S[\tilde{E}, f] := \int_S f^a \epsilon_{abc} e_n{}^b e_p{}^c \mathrm{d}x^n \wedge \mathrm{d}x^p = \int_S f_a \epsilon_{mnp} \tilde{E}^{ma} \mathrm{d}x^n \wedge \mathrm{d}x^p \,. \tag{2.41}$$

We note that the standard notation for this variable in the LQG literature is $E[S, f]$, but we prefer the one above because it is in parallel with the notation for the holonomy itself. Both $F_S[\tilde{E}, f]$ and $F^a_S[\tilde{E}]$ are *distributional* in the sense that they are supported on a two-dimensional submanifold of Σ.

To compute the Poisson brackets between the new canonical variables introduced above, we consider a surface S and an edge e that 'pierces' S at the point P (if e does not intersect S, the bracket simply vanishes). We next subdivide this edge into three pieces as shown in figure 1: two subedges e_1 and e_2, with associated holonomies $h^\epsilon_{e_1}$ and $h^\epsilon_{e_2}$, which touch S only in the limit $\epsilon \to 0$, and a third 'infinitesimal' edge $e(\epsilon)$ intersecting S, for which the path ordered exponential can be approximated by the linear term. Then

$$\begin{aligned}
&\Big\{ (h_e[A])_{\alpha\beta},\, F_S[\tilde{E}, f] \Big\} \\
&= \lim_{\epsilon \to 0} \left[\big(h^\epsilon_{e_1}[A]\big)_{\alpha\gamma} \left\{ \int_{e(\epsilon)} \mathrm{d}x^m A_m{}^a(\mathbf{x}) \tau^a_{\gamma\delta} \;,\; \int_S \mathrm{d}y^n \wedge \mathrm{d}y^p \epsilon_{npq} f_b(\mathbf{y}) \tilde{E}^{bq}(\mathbf{y}) \right\} \big(h^\epsilon_{e_2}[A]\big)_{\delta\beta} \right] \\
&= \iota(e,S) \gamma\, f_a(P) \big(h_{e_1}[A]\, \tau^a\, h_{e_2}[A] \big)_{\alpha\beta} \,.
\end{aligned} \tag{2.42}$$

Here, the intersection number

$$\iota(e,S) := \int_e \mathrm{d}x^m \int_S \mathrm{d}y^n \mathrm{d}y^p\, \epsilon_{mnp}\, \delta^{(3)}(\mathbf{x}, \mathbf{y}) = \pm 1 \;\text{ or }\; 0 \,, \tag{2.43}$$

encodes the information on how e intersects S in a coordinate independent way. The integral is equal to ± 1, depending on the orientation of e and S, if e intersects S transversally. When

the edge e intersects or touches S tangentially, a little care must be exercised; one then finds that $\iota(e,S) = 0$, and the above bracket vanishes [68]. The fact that $h_e[A]$ and $F_S[A,f]$ are, respectively, supported on one-dimensional and two-dimensional subsets of Σ is thus precisely what is required to perform the integral over the three-dimensional δ-function. Evidently, the integral is ill-defined when an entire segment of e lies within S; for a discussion of how to deal with this difficulty, see [68].

3 Quantisation: kinematics

Having determined the classical canonical variables, one would now like to promote them to quantum operators obeying the appropriate commutation relations. The essential assumption of LQG is that this quantisation should take place at the level of the bounded hermitean operators $h_e[A]$ rather than the connection A itself. This is analogous to ordinary quantum mechanics, when one replaces the Heisenberg operators x and p by Weyl operators e^{ix} and e^{ip}; the spin network representation actually uses the analog of a hybrid formulation with x and e^{ip}. The Stone-von Neumann theorem [69, 70, 71] is usually invoked to argue that it makes no difference whether one quantises the Heisenberg or the Weyl algebra, *i.e.* that these quantisations are equivalent. The theorem does require, however, that the representations which are used are 'weakly continuous'. In the case of ordinary quantum mechanics, for example, this means that matrix elements of the operators corresponding to $e^{i\alpha x}$ and $e^{i\beta p}$ are smooth functions of the parameters α and β. In LQG the representations of operators do not satisfy this requirement.[7]

The failure of operators to be weakly continuous can be traced back to the very special choice of the scalar product (3.7) below, which LQG employs to define its kinematical Hilbert space $\mathcal{H}_{\rm kin}$. This Hilbert space does not admit a countable basis, hence is *non-separable*, because the set of all spin network graphs in Σ is uncountable, and non-coincident spin networks are orthogonal w.r.t. (3.7). Therefore, any operation (such as a diffeomorphism) which moves around graphs continuously (w.r.t. to some background metric) corresponds to an uncountable sequence of mutually orthogonal states in $\mathcal{H}_{\rm kin}$. That is, no matter how 'small' the deformation of the graph in Σ, the associated elements of $\mathcal{H}_{\rm kin}$ always remain a finite distance apart, and consequently, the continuous motion in 'real space' gets mapped to a highly discontinuous one in $\mathcal{H}_{\rm kin}$. Although unusual, and perhaps counter-intuitive, as they are, these properties provide a cornerstone for the hopes that LQG can overcome the seemingly unsurmountable problems of conventional geometrodynamics: if the representations used in LQG were equivalent to the ones of geometrodynamics, there would be no reason to expect LQG not to end up in the same quandary.

It is perhaps also instructive to contrast the LQG approach with the standard lattice approach to field theory. In the lattice approach, all quantities depend explicitly on the lattice spacing (*i.e.* the regulating parameter). In the limit in which the lattice spacing is taken to zero, one recovers the continuum results (all expectation values are smooth functions of the regulating parameter). In the LQG approach the 'discretuum' is instead built in by the very construction of the scalar product, rather than by introducing a regulating parameter. While this radical modification of short-distance physics has certain appealing properties, it makes it hard to recover long-distance physics and the usual notion of continuity.

In short, one thus implements the quantisation not by the replacement (2.29) but rather by

[7] Strictly speaking, the Stone-von Neumann theorem also requires that the number of degrees of freedom is finite, which is also not true in canonical gravity.

promoting the above Poisson bracket (2.42) to a quantum commutator:

$$\Big[(h_e[A])_{\alpha\beta}\,,\,\hat{F}_S[\tilde{E},f]\Big] = i\hbar l_P^2\gamma\,\iota(e,S)\,f^a(P)\big(h_{e_1}[A]\,\tau_a\,h_{e_2}[A]\big)_{\alpha\beta} \tag{3.1}$$

or equivalently,

$$\Big[(h_e[A])_{\alpha\beta}\,,\,\hat{F}^a_S[\tilde{E}]\Big] = i\hbar l_P^2\gamma\,\iota(e,S)\,\big(h_{e_1}[A]\tau^a\,h_{e_2}[A]\big)_{\alpha\beta}\,. \tag{3.2}$$

On the spin network representation the holonomies $h_e[A]$ will be represented as multiplication operators; the action of the canonically conjugate operators will be explained below. The inequivalence of LQG quantisation with Fock-space quantisation arises through a special scalar product, to be discussed below.

3.1 The Hilbert space of spin networks

After defining the basic variables in which the theory should be quantised, the next step is to choose a Hilbert space in which the operators act. Starting from this space, one should construct a Hilbert space of physical states, *i.e.* space of states for which all constraints hold. The initial Hilbert space of LQG is the space of spin networks. While the Gauss constraint is easily solved in this space it turns out that a solution of the diffeomorphism constraints lies *outside* this 'naive' initial space, and one is forced to introduce a larger space.

The intuitive idea behind spin networks is that the geometry at the Planck scale is foam-like. The metric is excited only on one-dimensional sub-manifolds, which are called spin networks. Geometries which look smooth at large scales are supposed to arise only from complicated spin network states with many edges. In order to find the Hilbert space of these objects, one has to find a basis of wave functions over the configuration space, which associate a complex number to each and every configuration of the gauge connection. LQG makes use of wave functions which have singular support in the sense that they only probe the gauge connection on one-dimensional networks embedded in the three-dimensional spatial hypersurface Σ. This three dimensional 'reference space', or 'carrier space' of the spin networks, does not carry any physical metric. However, LQG makes occasional use of fiducial background coordinates and background metrics for certain intermediate steps in the construction, on which physical quantities are not supposed to depend. We will here avoid as far as possible using such background data. Let us also emphasise that the 'discreteness' of the spin networks does not correspond to a naive discretisation of space. Rather, the underlying continuum, on which the spin networks 'float', namely the spatial manifold Σ, is still present.

By definition, each network is a (not necessarily connected) graph Γ embedded in Σ and consisting of *finitely many* edges $e_i \in \Gamma$ and vertices $v \in \Gamma$. The edges are connected at the vertices. Each edge e carries a holonomy $h_e[A]$ of the gauge connection A (this connection does not have to be smooth on the edge). The wave function on the spin network over the graph Γ can be written as

$$\Psi_{\Gamma,\psi}[A] = \psi\big(h_{e_1}[A],\,h_{e_2}[A],\,\dots\big)\,, \tag{3.3}$$

where the ψ is some function of the basic holonomies associated to the edges $e \in \Gamma$. If, in addition, the wave function ψ is invariant under arbitrary SU(2) gauge transformations it satisfies the *Gauss constraint*, and vice versa. A gauge invariant function ψ thus takes care of joining the collection of holonomies into an SU(2)-invariant complex number by contracting all SU(2)-indices of the holonomies with invariant tensors 'located' at the vertices v, see figure 2. The basic building blocks of the spin network wave functions are therefore expressions of the following type. A three-valent vertex connects three edges according to

$$\psi[\text{fig.1a}] = \Big(\rho_{j_1}(h_{e_1}[A])\Big)_{\alpha_1\beta_1}\,\Big(\rho_{j_2}(h_{e_2}[A])\Big)_{\alpha_2\beta_2}\,\Big(\rho_{j_3}(h_{e_3}[A])\Big)_{\alpha_3\beta_3}\,C^{j_1j_2j_3}_{\beta_1\beta_2\beta_3}\dots\,, \tag{3.4}$$

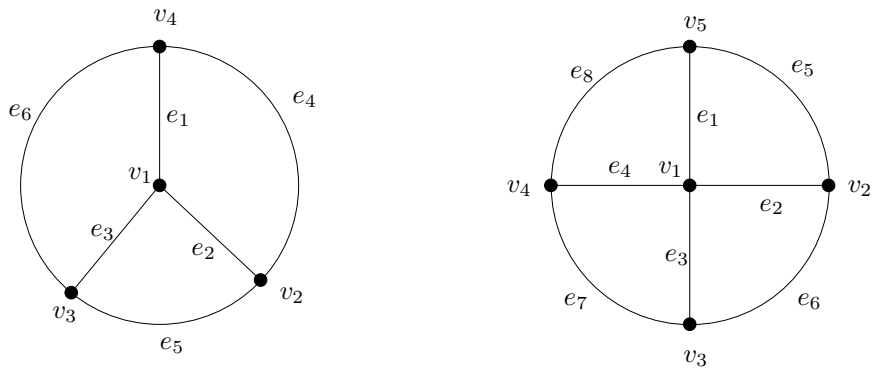

Figure 2: Examples of spin network states. For the 3-valent vertices on the left, the three incoming edges at each vertex are connected by a Clebsch-Gordan coefficient. For the 4-valent vertex on the right, one has to decide on a given way to construct a higher-order invariant tensor from two Clebsch-Gordan coefficients.

where dots represent the remainder of the graph. For a four-valent vertex, one first has to decide on how to pair the edges into two groups of two; one such choice leads to *e.g.*

$$\psi[\text{fig.1b}] = \Big(\rho_{j_1}(h_{e_1}[A]\Big)_{\alpha_1\beta_1} \Big(\rho_{j_2}(h_{e_2}[A])\Big)_{\alpha_2\beta_2} \Big(\rho_{j_3}(h_{e_3}[A])\Big)_{\alpha_3\beta_3} \Big(\rho_{j_4}(h_{e_4}[A])\Big)_{\alpha_4\beta_4} \times C^{j_1 j_2 k}_{\beta_1\beta_2\beta} C^{j_3 j_4 k}_{\beta_3\beta_4\beta} \cdots . \quad (3.5)$$

In these equations $C^{j_1 j_2 k}_{\beta_1\beta_2\beta}$ are intertwiners (Clebsch-Gordan coefficients). In (3.5) the intermediate spin k can be freely chosen in accordance with the standard rules for the vector addition of angular momenta: $|j_1 - j_2| \leq k \leq j_1 + j_2$. In other words, we can graphically represent the 4-valent vertex by splitting it into two 'virtual' 3-valent vertices and adding a 'virtual' edge, carrying angular momentum k (see figure 3). The same wave function can be re-expressed by performing this split in a different 'channel' by means of recoupling relations for the Clebsch Gordan coefficients (see for instance [72])

$$C^{j_1 j_2 k}_{\alpha_1\alpha_2\beta} C^{j_3 j_4 k}_{\alpha_3\alpha_4\beta} = \sum_m \sqrt{2k+1}\sqrt{2m+1} \begin{Bmatrix} j_1 & j_2 & k \\ j_3 & j_4 & m \end{Bmatrix} C^{j_1 j_3 m}_{\alpha_1\alpha_2\beta} C^{j_2 j_4 m}_{\alpha_3\alpha_4\beta}\,, \quad (3.6)$$

where the object with curly brackets is the Wigner $6j$-symbol. While the contraction is obviously unique if only two or three edges meet at a vertex, there may be more and independent choices for vertices of valence four and higher, depending on the way in which the edges are connected with the Clebsch-Gordan coefficients. For this reason, any given spin network will in general admit several independent wave functions of the above type. To make the dependence on the spins j_e associated with the edges and the intertwiners C_v associated with the vertices explicit, we will occasionally designate wave functionals such as (3.3) also by $|\Gamma, \{j\}, \{C\}\rangle$. The space of finite linear combinations of such states is denoted by $\mathcal{S}$ (for 'spin networks'). Although three-valent spin networks are obviously simplest, we will see later (see section 3.4.2) that higher valence is even generic because three-valent networks correspond to 'zero volume', and hence are deemed to be not of much interest.

The wave functionals (3.3) are called *cylindrical*, because they probe the connection A only 'on a set of measure zero' (like the δ-function does for ordinary functions). Similar 'cylindrical functionals' were formerly used in constructive quantum field theory in order to rigorously define the functional measures for free and certain interacting models of quantum field theory as limits

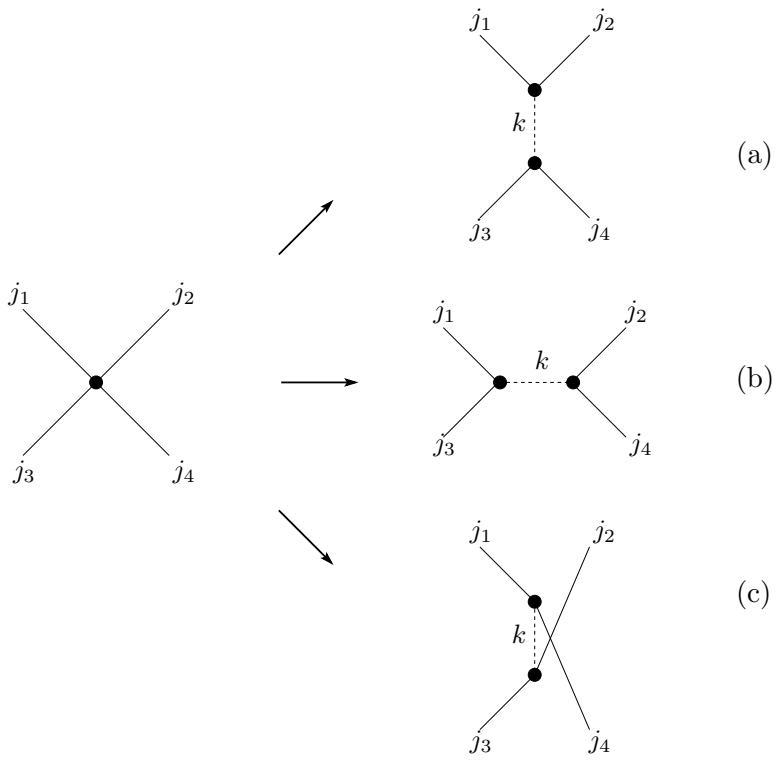

Figure 3: A four-valent vertex is defined by a particular way of connecting three-valent vertices. The spin k has to satisfy the triangle inequalities, but is otherwise arbitrary.

of finite dimensional integration measures [67]. The space $\mathcal{S}$ spanned by finite linear combinations of such cylindrical functions over all possible graphs is the starting point for the construction of the Hilbert space of spin networks. Obviously, the product of two cylindrical functions supported on the same or different spin networks is again a cylindrical function.

To complete the definition of the space of spin network states, we must introduce a suitable scalar product. In LQG this is not the standard scalar product induced by a Fock space representation (see the following section for more on this); instead, the scalar product of two cylindrical functions $\Psi_{\Gamma,\{j\},\{C\}}[A]$ and $\Psi_{\Gamma',\{j'\},\{C'\}}[A]$ is defined as

$$\begin{aligned}&\left\langle \Psi_{\Gamma,\{j\},\{C\}} \,\middle|\, \Psi'_{\Gamma',\{j'\},\{C'\}} \right\rangle \\ &\qquad = \begin{cases} 0 & \text{if } \Gamma \neq \Gamma' \,, \\ \displaystyle\int \prod_{e_i \in \Gamma} \mathrm{d}h_{e_i} \;\; \bar{\psi}_{\Gamma,\{j\},\{C\}}\Big(h_{e_1}, \dots\Big)\, \psi'_{\Gamma',\{j'\},\{C'\}}\Big(h_{e_1}, \dots\Big) & \text{if } \Gamma = \Gamma' \,, \end{cases} \end{aligned} \tag{3.7}$$

where the integrals $\int \mathrm{d}h_e$ are to be performed with the SU(2) Haar measure. Thus the inner product vanishes if the graphs Γ and Γ' do not coincide (even if they are 'very close to each other' in any given fiducial background metric). If Γ and Γ' coincide, the product may still vanish, depending on the choice of spins and intertwiners. At any rate, for $\Gamma = \Gamma'$, the inner product is given by integrating the product of $\bar{\psi}$ and ψ' over the holonomies at all edges, using the standard Haar measure. Because Γ was assumed to consist of *finitely* many edges, this is always a *finite-dimensional* integral, with one SU(2) integral for every edge in the graph. It is noteworthy that the above scalar product is *invariant under spatial diffeomorphisms*, even if the states Ψ_1 and Ψ_2 themselves are not, because the statement whether two graphs coincide or not

is diffeomorphism invariant. Whereas the original wave functionals $\Psi[A]$ probe the value of the connection and therefore also depend on the position of the graph, this information is 'lost' in the scalar product (3.7) which makes no more reference to the underlying space of connections or the 'shape' of the spin network graph.

One now defines the *kinematical Hilbert space* $\mathcal{H}_{\rm kin}$ as the completion of the space of gauge invariant spin network states $\mathcal{S}$ w.r.t. to this norm. $\mathcal{H}_{\rm kin}$ thus consists of all linear superpositions of spin network states $\Psi_n \equiv \Psi_{\Gamma_n,\{j_n\},\{C_n\}}$ such that they have finite norm,

$$\Psi = \sum_{n=1}^{\infty} a_n \Psi_n \quad , \qquad ||\Psi||^2 = \sum_{n=1}^{\infty} |a_n|^2 \, ||\Psi_n||^2 < \infty \tag{3.8}$$

where the norm $||\Psi_n||^2$ is defined by (3.7). One of the distinctive features of the Hilbert space $\mathcal{H}_{\rm kin}$, which will play a key role in the further development of the theory, is its *non-separability*. This non-separability can be traced back to the existence of an underlying continuum, the spatial manifold Σ, and accounts for the difference between this scheme and the standard (Fock space) quantisation of gauge theories, see the following subsection. Although each Ψ_n in the above sum is associated to a *finite* graph Γ_n the expected number of edges need not be finite, because the sum

$$\langle \#(\text{edges}) \rangle \; \propto \; \sum_{n=1}^{\infty} |a_n|^2 \, L(\Gamma_n) ||\Psi_n||^2 \, , \tag{3.9}$$

with $L(\Gamma_n)$ the number of edges in Γ_n, can be made to diverge if $L(\Gamma_n)$ increases sufficiently rapidly with n, even if $\sum_n |a_n|^2 \, ||\Psi_n||^2 < \infty$. Idem for the expectation value for the number of vertices. Let us, however, caution readers already at this point that $\mathcal{H}_{\rm kin}$ is *not* the relevant Hilbert space for solving the quantum constraints in LQG; we will have more to say about this in the following sections.

A second unusual feature is the *ab initio* absence of negative norm states, despite the fact that at this stage the constraints have not even appeared yet. This is in stark contrast to the usual covariant Fock space quantisation of gauge theories, where negative norm states are unavoidable, and can only be eliminated (if they can be eliminated at all) by restricting the Hilbert space to the subspace of *physical states*, which by definition are annihilated by the constraints. Curiously, there does exist a reformulation of (free) second quantised Maxwell theory with similar properties [73]. Roughly speaking, this consists in (*i*) defining the scalar product to be identical with the standard Fock space one for gauge invariant expectation values (as is well known, in second quantised Maxwell theory this restriction indeed eliminates negative norm states), and (*ii*) by declaring gauge-variant scalar products to be zero. In this way one preserves Lorentz covariance (because a gauge variant state is never transformed into a gauge invariant one by a Lorentz transformation, and vice versa), and eliminates negative norm states, in such a way that *for observable quantities the results are identical with the usual Fock space quantisation.* The price one pays is the loss of weak continuity (because the expectation values for gauge variant expressions always vanish, no matter how close they are to gauge invariant ones) and the loss of separability of the Hilbert space. These are precisely the features encountered in the LQG quantisation procedure above. However, the crucial difference is that the prescription [73] is tailored to reproduce the standard Fock space quantisation in the observable sector, whereas LQG does not. Also, it is not known how to extend the prescription of [73] to full QED, or how to implement the procedure in a perturbative approach (with unphysical intermediate states in Feynman diagrams).

3.2 Comparison with the loop approach in gauge theories

The (Wilson) loop approach has been advocated in past in the context of gauge theories, as an attempt to non-perturbative quantisation. This approach eventually did not lead to the expected success, see [74, 75] for detailed reviews of the subject. Nevertheless LQG practitioners argue that their approach, despite its similarity to the loop approach in gauge theories, should work for gravity. Because this relates to one of the frequently asked questions about the LQG programme, we here briefly contrast some key features of the Wilson loops in gauge theory with the behaviour of spin networks in LQG. Not unexpectedly, the essential difference turns out to lie in the special scalar product (3.7) used in LQG.

The first difference between Wilson loops and spin networks is that, due to the product (3.7), any spin network is orthogonal to any other spin network except itself, including the trivial empty one (*i.e.* the vacuum). Hence,

$$\langle\, \Psi_{\Gamma,\{C\}}[A]\,\rangle = \langle\, \mathbf{1}\,|\,\Psi_{\Gamma,\{C\}}[A]\,\rangle = 0\,. \tag{3.10}$$

By contrast, in gauge theory, the expectation value of a single Wilson loop encodes key physical information: it determines the behaviour and properties of quark - anti-quark potentials, and can signal confinement of quarks. More specifically, in a perturbative approach we have

$$\langle\, W_{\mathcal{C}}\,\rangle \equiv \left\langle \exp\left[ie \oint_{\mathcal{C}} A_\mu \mathrm{d}x^\mu\right]\right\rangle = \exp\left[-\frac{ie}{2}\oint_{\mathcal{C}}\oint_{\mathcal{C}} \mathrm{d}x^\mu \mathrm{d}x^\nu\, \Delta_{\mu\nu}(x-y)\right]\,. \tag{3.11}$$

where $\Delta_{\mu\nu}$ is the gluon propagator. In conventional quantum field theory (QCD), this expression is divergent, in contrast to (3.10)) because of the self-interaction of the quarks and gluons propagating in the loop. The way in which it diverges generically depends on the shape of the loop. Nevertheless, the UV divergences in (3.11) can be consistently removed, for example by point splitting regularisation, and a well defined answer obtained. All loops can be regularised in this way (while it is only known how to renormalise, *smooth* and *non-self-intersecting* loops [74]). A better, and entirely non-perturbative scheme to calculate the Wilson loops and other quantities of interest is, of course, provided by lattice gauge theory [76, 77, 78].

In the case of correlation functions of several Wilson loops, no new types of divergences appear. For two non-intersecting loops $\mathcal{C}$ and $\mathcal{C}'$ 'self energy divergences' can be removed by defining the 'connected correlator'

$$\langle W_{\mathcal{C}} W_{\mathcal{C}'}\rangle_{\mathrm{conn}} := \langle W_{\mathcal{C}} W_{\mathcal{C}'}\rangle - \langle W_{\mathcal{C}}\rangle\langle W_{\mathcal{C}'}\rangle = \text{finite}\,. \tag{3.12}$$

which can be thought of as the analog of the scalar product of two single loops in LQG. In first approximation we then obtain

$$\langle\, W_{\mathcal{C}_1}\,|W_{\mathcal{C}_2}\rangle_{\mathrm{conn}} = \exp\left[-\frac{ie}{2}\oint_{\mathcal{C}_1} \mathrm{d}x^\mu \oint_{\mathcal{C}_2} \mathrm{d}y^\nu\, \Delta_{\mu\nu}(x-y)\right]\,. \tag{3.13}$$

The main point about this result is that it depends continuously on the shape of the loops $\mathcal{C}_1$ and $\mathcal{C}_2$ – unlike the scalar product or correlator of two spin networks (3.7), which is non-zero if and only if the two networks completely overlap. Indeed, different shapes of Wilson loops in gauge theory are inequivalent, such that the relative change can be expressed via the Yang-Mills fields strength and the 'non-abelian Stokes' theorem'. The value of the Wilson loop is invariant under continuous deformations only for vanishing field strength (as in a topological theory, such as Chern-Simons gauge theory). In LQG on the other hand, the physical states (for which we have to enlarge the space $\mathcal{S}$, as we said) are supposed to be diffeomorphism invariant. Hence there is no physical information whatsoever in the shape of the loop, since two networks of different shape

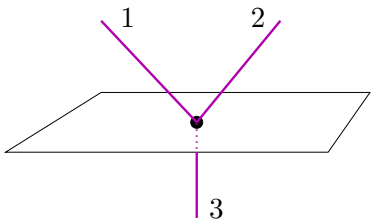

Figure 4: A flux operator intersecting a three-valent node. The action is given by equation (3.15).

but identical topology can always be related by a suitable diffeomorphism. This equivalence holds independently of the 'value' of the Ashtekar field strength in this state.

When matter is included in the LQG formalism, consistency would seem to require that standard gauge fields must be treated analogously to the Ashtekar connection, *i.e.* with holonomies associated to edges, and a scalar product similar to (3.7), leading to a non-separable kinematical Hilbert space for the matter sector as well. It is not known how (and whether) this formalism can recover known results such as the ones sketched above, and in particular the shape dependence of the Wilson loops, and we do not know of any attempts in this direction. At the very least, this will require a sophisticated analog of the 'shadow states' introduced in [21]. For kinematical states (which do not satisfy the constraints) progress along these lines has been made in [22, 23, 24, 25, 26, 27].

3.3 The action of elementary operators on a spin network

We would now like to implement the canonical variables as operators on spin network wave functions in such a way that the canonical commutation relations (3.1) and (3.2) are satisfied. This will provide us with the 'building blocks' that will allow us to define the action of the basic kinematical operators (area and volume operators) and the Hamiltonian constraint operator on any spin network wave function. The elementary holonomy and flux operators defined in (2.35) and (2.41) act in a rather simple way on the spin network states. Namely, the holonomy $(h_e[A])_{\alpha\beta}$ simply acts as a multiplication operator, and thus adds the edge e to an existing network Γ. This edge may be disconnected from or 'overlaid' on Γ; usually it will appear in some gauge invariant combination, whose action on the spin network wave function produces another such wave function. The action of the flux $\hat{F}_S[\tilde{E}, f]$ on a network state is either zero (if the surface element S does not intersect the graph Γ anywhere), or otherwise cuts in two any edge which pierces S, and inserts a new intertwiner vertex τ_a into the spin network at the point of intersection. Consequently, on such a single edge (which is part of some spin network wave function $\Psi_{\Gamma,\{C\}}$) one has

$$\hat{F}^a_S[\tilde{E}] \left(\dots \big(h_e[A]\big)_{\alpha\beta} \dots \right) = 8\pi i l_P^2 \hbar\gamma \, \iota(e,S) \left(\dots \big(h_{e_1}[A]\, \tau^a \, h_{e_2}[A]\big)_{\alpha\beta} \dots \right). \tag{3.14}$$

where $e = e_1 \cup e_2$ and we have now omitted the symbol ρ_{j_e}. Dots indicate the remaining part of the spin network wave function.

If $\hat{F}^a_S[\tilde{E}]$ cuts the network at a vertex, the result depends on the choice for the intertwiner as well as the orientation of the surface S with respect to the edges. Let us illustrate this by considering the 3-valent vertex depicted in figure 4. The flux acts at the end of every edge that emanates from the node. Edges which are located above the surface S contribute with opposite sign from the ones that are located below, and for the example of the 3-valent vertex the net

result is

$$\begin{aligned}
\hat{F}^a_S[\tilde{E}]&\Big(C^{j_1 j_2 j_3}_{\alpha_1\alpha_2\alpha_3}\big(h_{e_1}[A]\big)_{\alpha_1\beta_1}\big(h_{e_2}[A]\big)_{\alpha_2\beta_2}\big(h_{e_3}[A]\big)_{\alpha_3\beta_3}\Big)\\
&= 8\pi l_P^2\hbar\gamma\Big(\big(\rho_{j_1}(\tau^a)\big)_{\alpha_1\gamma_1}C^{j_1 j_2 j_3}_{\gamma_1\alpha_2\alpha_3} + \big(\rho_{j_2}(\tau^a)\big)_{\alpha_2\gamma_2}C^{j_1 j_2 j_3}_{\alpha_1\gamma_2\alpha_3} - \big(\rho_{j_3}(\tau^a)\big)_{\alpha_3\gamma_3}C^{j_1 j_2 j_3}_{\alpha_1\alpha_2\gamma_3}\Big)\\
&\qquad\times\Big(\big(h_{e_1}[A]\big)_{\alpha_1\beta_1}\big(h_{e_2}[A]\big)_{\alpha_2\beta_2}\big(h_{e_3}[A]\big)_{\alpha_3\beta_3}\Big) \qquad (3.15)\\
&= -16\pi l_P^2\hbar\gamma\,\big(\rho_{j_3}(\tau^a)\big)_{\alpha_3\beta_3}C^{j_1 j_2 j_3}_{\alpha_1\alpha_2\beta_3}\Big(\big(h_{e_1}[A]\big)_{\alpha_1\beta_1}\big(h_{e_2}[A]\big)_{\alpha_2\beta_2}\big(h_{e_3}[A]\big)_{\alpha_3\beta_3}\Big)\,.
\end{aligned}$$

(the uncontracted indices $\beta_1, \beta_2, \beta_3$ connect to the remainder of the spin network wave function). In the last step, we have used the invariance property of the Clebsch-Gordan coefficients,

$$\Big(\rho_{j_1}(\tau^a)\Big)_{\alpha_1\beta_1}C^{j_1\,j_2\,j_3}_{\beta_1\alpha_2\alpha_3} + \Big(\rho_{j_2}(\tau^a)\Big)_{\alpha_2\beta_2}C^{j_1\,j_2\,j_3}_{\alpha_1\beta_2\alpha_3} + \Big(\rho_{j_3}(\tau^a)\Big)_{\alpha_3\beta_3}C^{j_1\,j_2\,j_3}_{\alpha_1\alpha_2\beta_3} = 0\,. \qquad (3.16)$$

Likewise, when the node has higher valence, the action of the flux operator $\hat{F}^a_S[\tilde{E}]$ inserts factors $\pm\rho_{j_e}(\tau^a)$ in the appropriate place in the spin network wave function, with a '+' for edges e 'above' S, and a '−' for the edges 'below' S. This action is sometimes symbolically represented in terms of angular momentum operators $\hat{J}^a_{(e,v)}$. Consequently, the flux operator can be re-expressed as a sum of angular momentum operators, viz.

$$\hat{F}^a_S[\tilde{E}] = \sum_e \sigma(e,S)\hat{J}^a_{(e,v)} \qquad (3.17)$$

where the sum runs over all edges entering the vertex v, and the signs $\sigma(e,S) = \pm 1$ are determined by the position of e in relation to S. When the signs are all '+', the sum vanishes for gauge invariant wave functionals by angular momentum conservation, *i.e.* Gauss' law.

3.4 The kinematical operators: area and volume operators

Two important *kinematical operators* in quantum gravity are the volume and the area operators. The area operator is relevant in the application of LQG to black hole physics, while the volume operator is essential for the very definition and the implementation of the scalar (Hamiltonian) constraint. It is important to stress that neither of these operators is an observable *à la* Dirac, in the sense that neither commutes with the canonical constraints (classically or quantum mechanically). Their supposed physical relevance is based on the general belief that proper Dirac observables corresponding to area and volume should exist, provided suitable matter degrees of freedom are included. The reason is that, for an operational definition of area and volume, one needs 'measuring rod fields' in the same way that the operational definition of 'time' in quantum gravity requires a 'clock field', and that the true observables are appropriate combinations of gravitational and matter fields (which then would commute with the constraints). However, these issues are by no means conclusively settled (for recent work on this topic see [79, 80]), and we are not aware of any concrete realisation of this intuitive idea.

Setting aside such conceptual worries, the main strategy in constructing the area and volume operators is to first write the classical expressions for these quantities using the basic variables (holonomies and electric field), and then promote those into operators. To do this one first has to regularise the classical expressions by a discretisation of the integrals [81, 82]. It will turn out that the ambiguity in the way one does this is irrelevant for the area operator, while it does affect the result for the volume operator.

We should point out that the area and volume operators appear to be of no particular relevance in conventional geometrodynamics. There, they would simply be represented by multiplication operators (defining a 'length operator' would also be straightforward). As long as these act on smooth wave functionals, there would be no problem, and issues of regularisation and renormalisation, like the ones we are about to discuss, would not even arise. On the other hand, on more singular 'distributional' functionals, these operators would not be well defined, as they depend non-linearly on the basic metric variables. Whether it is possible or not to properly define these operators in geometrodynamics therefore depends on what the 'good' and the 'bad' wave functionals are. But this question is impossible to answer in the absence of a suitable scalar product and a Hilbert space of wave functionals (see e.g. [19] for a discussion of these and other difficulties in geometrodynamics).

3.4.1 The area operator

A two-dimensional surface $S \subset \Sigma$ possesses a unit normal vector field $\vec{N}$ and its classical area is given by the determinant of the induced metric $h_{mn} = g_{mn} - N_m N_n$. To compute the area, it is however simpler to work directly with the area element (2.38) (cf. subsection 2.4), so that

$$A_S[g] = \int_S \sqrt{\mathrm{d}F^a \cdot \mathrm{d}F^a}\,. \tag{3.18}$$

To convert this into an expression that can act on the spin network wave function, we subdivide the surface into infinitesimally small surfaces S_I (with $I = 1, \ldots, N$), such that $S = \cup_I S_I$. Next, we approximate the area by a Riemann sum (which, of course, converges for well-behaved surfaces S), using

$$\int_{S_I} \sqrt{\mathrm{d}F^a \cdot \mathrm{d}F^a} \approx \sqrt{F^a_{S_I}[\tilde{E}]\, F^a_{S_I}[\tilde{E}]}\,. \tag{3.19}$$

Therefore

$$A_S[\tilde{E}^a_m] = \lim_{N\to\infty} \sum_{I=1}^{N} \sqrt{F^a_{S_I}[\tilde{E}]\, F^a_{S_I}[\tilde{E}]}\,. \tag{3.20}$$

The Riemann sum on the right-hand side of (3.20) can now, for fixed N, be promoted unambiguously to an operator. Namely, when acting on an arbitrary spin network wave function, the area operator will receive a contribution each time an edge of the spin network pierces an *elementary* surface S_I in the sum (3.20). More precisely, one can use equation (3.14) to derive (always in the appropriate representations $\rho_{j_{e_i}}$ and assuming $\iota(e,S) = \pm 1$)

$$\begin{aligned} \hat{F}^a_{S_I}[\tilde{E}]\, \hat{F}^a_{S_I}[\tilde{E}] \;\; \big(h_e[A]\big)_{\alpha\beta} &= i^2 (8\pi l_P^2 \hbar\gamma)^2 \left(h_{e_1}[A]\, \tau_a\, \tau^a\, h_{e_2}[A]\right)_{\alpha\beta} \\ &= (8\pi l_P^2 \hbar\gamma)^2 \left(j_e(j_e+1)\right) \big(h_e[A]\big)_{\alpha\beta}\,. \end{aligned} \tag{3.21}$$

Note also that if the the surface S_I had been pierced by two edges (carrying representations j_1 and j_2) instead of just one, the previous expression would generalise to

$$\begin{aligned} &\hat{F}^a_{S_I}[\tilde{E}]\, \hat{F}^a_{S_I}[\tilde{E}] \;\; \big(h_{e_1}[A]\big)_{\alpha\beta} \big(h_{e_2}[A]\big)_{\gamma\delta} \\ &\qquad = (8\pi l_P^2 \hbar\gamma)^2 \Big(j_{e_1}(j_{e_1}+1) + j_{e_2}(j_{e_2}+1)\Big) \big(h_{e_1}[A]\big)_{\alpha\beta} \big(h_{e_2}[A]\big)_{\gamma\delta}\,. \end{aligned} \tag{3.22}$$

Hence if one applies the operator (3.20) to wave function associated with a fixed graph Γ with $L(\Gamma)$ edges, and refines it in such a way that each elementary surface S_I is pierced by only *one*

edge of the network, one obtains

$$\hat{A}_S\Psi = 8\pi l_p^2\hbar\gamma \sum_{p=1}^{L(\Gamma)} \sqrt{j_p(j_p+1)}\,\Psi\,. \tag{3.23}$$

It is obvious that further refinement of the area operator then does not change the result. It is also clear that the final result, therefore, does not depend on the shape of the elementary area cells. From (3.23), one sees that all spin networks are eigenstates of the area operator (when the S_I intersect only edges, see below). Due to the 'discreteness' of the spin network states, the spectrum of the area operator is also discrete. The quantisation scale is set not only by the Planck length l_p^2, but also depends on the parameter γ.

We have so far assumed that the surface elements S_I in the above computation 'meet' only the edges (hence insert a bivalent vertex into the spin network), but not the vertices of the spin network. Making use of the general result (3.17) we can also evaluate the area operator when the surface elements S_I intersect a vertex. In this case, the result depends on the particular intertwiner at that node and on the orientation of the surface element. Let us illustrate this on a four-valent vertex, with intertwiner $C^{j_1j_2k}_{\alpha_1\alpha_2\beta}C^{j_3j_4k}_{\alpha_3\alpha_4\beta}$ (see figure 5). Consider first the situation in which the edges 1 and 2 are both located on the same side of the surface element and the same is true for edges 3 and 4 (figure 5a). In this case, one can use the invariance of the Clebsch-Gordan coefficients (3.16) to move around the inserted τ-matrices, and re-write the action of the area operator in terms of a 'virtual' edge. For our example, this yields the result

$$\hat{A}_S\,|\text{figure 5a}\rangle = 8\pi l_p^2\hbar\gamma\,\sqrt{2\,k(k+1)}\,|\text{figure 5a}\rangle\,. \tag{3.24}$$

Observe the relative factor of $\sqrt{2}$ with respect to (3.23): the result for a virtual edge is not the same as the one for a real edge. The four-valent node is thus an eigenstate of this area operator. Next, consider the situation in which the surface element is oriented as in figure 5b. In this case, one first has to rewrite the product of Clebsch-Gordan coefficients by means of the recoupling relation (3.6). Independently of the normalisation of the $6j$-symbol, the important thing to note is that the area operator will *change* the relative coefficients of the various terms in the sum (3.6), picking up a contribution $\propto \sqrt{m(m+1)}$ for each admissible value of the angular momentum $|j_1 - j_3| \leq m \leq j_1 + j_3$ in the new channel $1+3 \to 2+4$. Hence, the four-valent vertex is generally *not* an eigenstate of the area operator associated to the surface of figure 5b. For higher valences, the number of different contributions increases rapidly with the number of accessible channels. Therefore, the area operator in general does *not* act diagonally on an arbitrary spin network state.

We conclude this subsection with three remarks. Although infinite spin networks have been excluded *by fiat*, it is nonetheless instructive to see what would happen if we did include them. In this case it may not be possible to achieve the refinement of one edge per one cell in a *finite* number of steps. To see this, consider the example of a strongly fluctuating network, of 'shape' $z = 0$, $y = \sin(1/x)$ in the fiducial background coordinates. It is clear that if one considers the area of the $y = 0$ surface, then no matter how small the size of cells, the cell around zero will always be pierced by an *infinite* number of edges, and hence the area operator would not be well defined. Because classical space(-time?) presumably emerges in the limit of infinite spin networks, the exclusion of such 'ill-behaved' networks really amounts to a prescription for the order in which limits are to be taken, when they would otherwise produce an ambiguous answer.

Secondly, the area operator is unbounded, and not every state in $\mathcal{H}_{\text{kin}}$ possesses finite area (whereas every element of $\mathcal{S}$ has finite area). This can happen even for compact spatial manifold Σ, and the wave function can then be thought of as associated to an 'infinitely crumpled surface'

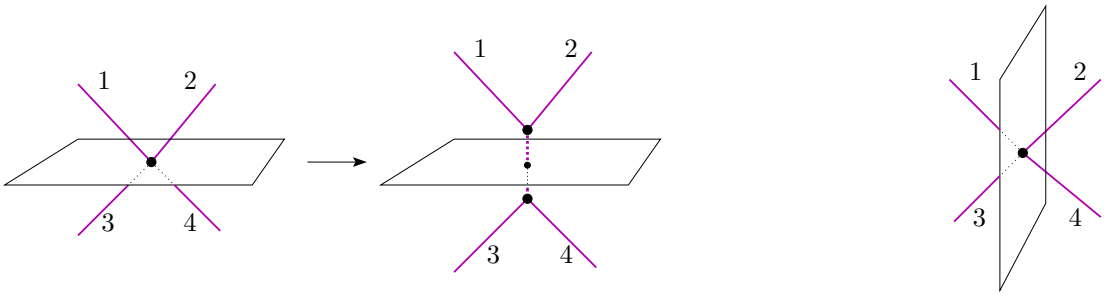

Figure 5: The action of the area operator on a node with intertwiner $C^{j_1 j_2 k}_{\alpha_1\alpha_2\beta} C^{j_3 j_4 k}_{\alpha_3\alpha_4\beta}$. In the figure on the left, the location of the edges with respect to the surface is such that the invariance of the Clebsch-Gordan coefficients (3.16) can be used to evaluate the action of the area operator. The result can be written in terms of a "virtual" edge. In the figure on the right, however, the edges are located in a different way with respect to the surface. The invariance property (3.16) does not apply, one has to use the recoupling relation (3.6), and the spin network state is therefore not an eigenstate of the corresponding area operator.

in Σ. To see this more explicitly, consider a sequence of spin network states Ψ_n with edges labelled by $p = 1, \ldots, L(\Gamma_n)$, and spins $j_p^{(n)}$; then we have

$$\hat{A}_S \left(\sum_{n=1}^{\infty} a_n \Psi_n \right) = 8\pi l_p^2 \hbar\gamma \sum_{n=1}^{\infty} \left(\sum_{p=1}^{L(\Gamma_n)} \sqrt{j_p^{(n)}(j_p^{(n)}+1)}\, a_n \Psi_n \right) . \tag{3.25}$$

The expectation value of $\hat{A}_S$ can be made to diverge by letting the spins $\{j_p^{(n)}\}$ grow sufficiently fast with n, even if $\sum_n |a_n|^2 ||\Psi_n||^2 < \infty$.

Thirdly, the formula (3.23) plays an important role in the microscopic explanation of the Bekenstein-Hawking entropy of black holes [34, 35], which is now reduced to a counting problem. Ideally, fitting the solution of this problem to the known factor $\frac{1}{4}$ relating entropy and area of the horizon would imply a prediction for the value of the parameter γ. Unfortunately, the results obtained so far disagree; for instance, [36] obtains $\gamma = \frac{\ln 2}{\pi\sqrt{3}}$, whereas a more recent calculation [83, 84] gives $\gamma = 0.23753295...$. We suspect that the final word has not yet been spoken on this subject [85].

3.4.2 The volume operator

The construction of the volume operator is less 'clean cut' than that of the area operator, because it is fraught with ambiguities, which can only be eliminated by making certain choices and by 'averaging', see [82] for a comprehensive discussion. In a first approximation one might try to use the functional differential operator (2.31), but its direct application would lead to to singular factors $\delta^{(3)}(\mathbf{x}, \mathbf{x})$. The 'detour' in the construction to be described, in particular the smearing over the surfaces S_I^a in (3.30) below, is necessary mainly in order to eliminate these singular factors. One starts with the classical expression for the volume of a three-dimensional region $\Omega \subset \Sigma$,

$$V(\Omega) = \int_\Omega \mathrm{d}^3x \sqrt{\left| \frac{1}{3!} \epsilon_{abc} \epsilon^{mnp} \tilde{E}^a_m \tilde{E}^b_n \tilde{E}^c_p \right|} . \tag{3.26}$$

Just as with the area operator, one partitions Ω into small cells $\Omega = \cup_I \Omega_I$, so that the integral can be replaced with a Riemann sum. To express the volume operator in terms of canonical

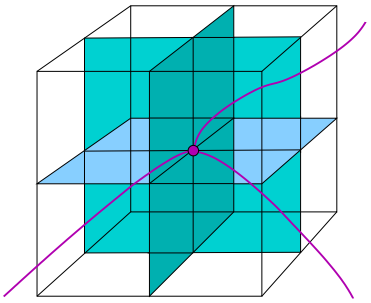

Figure 6: An elementary cell in the Riemann sum (3.29), together with the three surfaces on which the fluxes are evaluated.

quantities, we rewrite the volume element $\mathrm{d}V = \mathrm{d}\theta^1 \wedge \mathrm{d}\theta^2 \wedge \mathrm{d}\theta^3$ in terms of the area elements $\mathrm{d}F^1 = \mathrm{d}\theta^2 \wedge \mathrm{d}\theta^3$, ... as

$$\mathrm{d}V = \sqrt{|\mathrm{d}F^1 \cdot \mathrm{d}F^2 \cdot \mathrm{d}F^3|}\,. \tag{3.27}$$

In order to express this volume element in terms of the canonical quantities introduced before, we next approximate the area elements $\mathrm{d}F^a$ by the small but finite area operators $F^a_S[\tilde{E}]$, such that the volume is obtained as the limit of a Riemann sum

$$V(\Omega) = \lim_{N\to\infty} \sum_{I=1}^{N} \sqrt{\left|\frac{1}{3!}\epsilon_{abc} F^a_{S^1_I}[\tilde{E}]\, F^b_{S^2_I}[\tilde{E}]\, F^c_{S^3_I}[\tilde{E}]\right|}\,. \tag{3.28}$$

Now, in order to make sense of this expression, we must for each cell Ω_I choose three small non-coincident surfaces S^a_I, which subdivide Ω_I, as shown in figure 6. This should be done in such a way that the r.h.s. of (3.28) reproduces the correct classical value. For instance, one can choose a point inside each cube Ω_I, then connect these points by lines and 'fill in' the faces. In each cell Ω_I we then have three lines labelled by $a = 1, 2, 3$; the surface S^a_I is then the one that is traversed by the a-th line. With this choice it can be shown that the result is insensitive to small 'wigglings' of the surfaces, hence independent of the shape of S^a_I, and the above expression converges to the desired result [8].

Having constructed a sequence of refinements with the correct classical limit, we can now attempt to promote (3.28) to a quantum operator by invoking the known action of each factor $\hat{F}^a_{S^a_I}$ on a spin network wave function, viz.

$$\hat{V}(\Omega) = \lim_{N\to\infty} \sum_{I=1}^{N} \hat{V}(\Omega_I) = \lim_{N\to\infty} \sum_{I=1}^{N} \sqrt{\left|\hat{q}^{(N)}_I(S^1_I, S^2_I, S^3_I)\right|}\,, \tag{3.29}$$

with the operator (= square of the local volume)

$$\hat{q}^{(N)}_I(S^1_I, S^2_I, S^3_I) := \frac{1}{3!}\epsilon_{abc} \hat{F}^a_{S^1_I}[\tilde{E}]\, \hat{F}^b_{S^2_I}[\tilde{E}]\, \hat{F}^c_{S^3_I}[\tilde{E}]\,. \tag{3.30}$$

Ignoring the square root for the moment, (3.30) is indeed well defined as an operator at this point. However, unlike the classical volume, the action of $\hat{q}^{(N)}_I$ on the spin network wave function depends on the way we partition Ω, and more specifically on the position of the faces S^a_I in relation to the edges and vertices. As a simple example of the inherent ambiguities consider a cube which is traversed by one edge e, but not containing any vertex of the network. If the surfaces S^a_I are

[8] Alternatively, one can partition Σ with other types of cells, with analogous results.

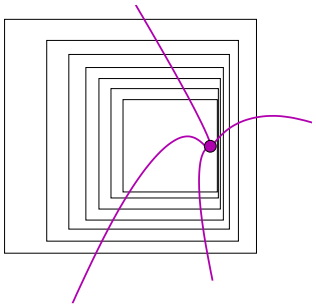

Figure 7: A series of refining steps for one cell in a Riemann sum, such that the indicated node is never at the centre of any of the boxes except in the strict $N \to \infty$ limit. A volume operator based on cells of this kind would always yield a vanishing result for any finite cell size.

positioned in such a way that the edge intersects each one of them with $\iota(e, S^a_I) \neq 0$ at three different points in Ω_I, the operator simply inserts the matrices τ^1, τ^2 and τ^3 at the appropriate places in $h_e[A]$. If the cell is sufficiently small, we can bring these matrices together, using $\tau^1\tau^2\tau^3 = i$, and thus pick up a contribution of order one. It is easy to construct sequences of refinements such that an arbitrarily large number of cells Ω_I gives such a contribution, no matter how small the cells are. In order to prevent catastrophic divergent sums from appearing on the r.h.s. of (3.29) one must therefore rule out such configurations 'by hand' and stipulate that each cell Ω_I may contain at most two such intersections with any edge, in which case one of the three factors in (3.30) will give zero.

With this proviso, the sum receives contributions only from those cells Ω_I which contain a vertex v. If the vertex v does not coincide with the intersection point of the three faces S^a_I, further refinement will yield more cells, which do not contribute by the above requirement. Hence, with sufficient refinement only those cells will remain and contribute for which the vertex v sits precisely at the intersection: $v = S^1_I \cap S^2_I \cap S^3_I$. This means that we must also exclude refinements such as in figure 7 for which $S^1_I \cap S^2_I \cap S^3_I$ 'misses' v by an amount that gets smaller with N, but never vanishes for $N < \infty$. With these assumptions, the resulting action of the operator $\hat{q}_I$ on the vertex v reduces to

$$\hat{q}_I = (8\pi l_P^2 \gamma)^3 \frac{1}{48} \sum_{e_1,e_2,e_3} \epsilon(e_1, e_2, e_3) \epsilon_{abc} \hat{J}^{(e_1)}_a \hat{J}^{(e_2)}_b \hat{J}^{(e_3)}_c \,, \tag{3.31}$$

where the sum runs over all possible triplets of edges in the vertex v, and the angular momentum operator $\hat{J}^{(e,v)}_a$ (defined in (3.17)) by definition inserts a matrix $\rho_{j_e}(\tau^a)$ at the vertex v into the spin network wave function. The coefficients $\epsilon(e_1, e_2, e_3)$ are orientation factors which are equal to ± 1. They depend on how the faces S^a_I cut across the vertex, and on which edges are 'above' and 'below' S^a_I. More precisely, in terms of the factors $\sigma(e, S)$ introduced in (3.17) they are given by

$$\epsilon(e_1, e_2, e_3) = \sigma(e_1, S^1_I)\, \sigma(e_2, S^2_I)\, \sigma(e_3, S^3_I) \tag{3.32}$$

and thus completely fixed by the relative positions of the edges and the surfaces. Let us also stress they are independent of any fiducial metric background structure, and purely combinatorial – hence invariant under diffeomorphisms. (3.32) is not yet the final answer, but let us first make the above formulas a little more concrete by working out the action of the above operator on a

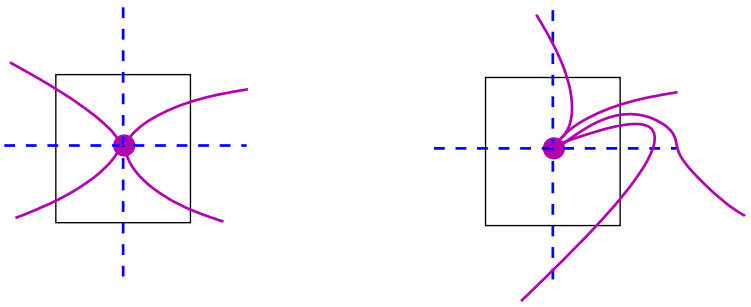

Figure 8: The vertex on the left can be deformed in such a way as to make the eigenvalue of the volume operator at this vertex vanish (prior to averaging).

4-valent vertex (3.5). Its net effect is the replacement of the intertwiner according to

$$C^{j_1\cdots j_4}_{\alpha_1\ldots\alpha_4} \to \text{const.} \sum_{\text{perm.}\{1,2,3,4\}} \epsilon(e_1,e_2,e_3)\,\epsilon_{abc} \times \Big(\rho_{j_{e_1}}(\tau^a)\Big)_{\alpha_1\beta_1} \Big(\rho_{j_{e_2}}(\tau^b)\Big)_{\alpha_2\beta_2} \Big(\rho_{j_{e_3}}(\tau^c)\Big)_{\alpha_3\beta_3} \delta_{\alpha_4\beta_4} C^{j_{e_1}\cdots j_{e_4}}_{\beta_1\ldots\beta_4}\,, \quad (3.33)$$

where the sum runs over all choices of three edges out of four. From this formula it is obvious that in general the operator $\hat{q}_I$ acts *non-diagonally* on a spin network wave function $|\Gamma,\{j\},\{C\}\rangle$, because it changes the intertwiners $\{C\}$, although it does not affect the spin network itself (*i.e.* the graph Γ). The operator vanishes when all the edges enter v through the same octant, in which case $\epsilon = +1$ for all choices, and

$$\epsilon_{abc}\big(\hat{J}^a_1\hat{J}^b_2\hat{J}^c_3 + \hat{J}^a_1\hat{J}^b_2\hat{J}^c_4 + \hat{J}^a_1\hat{J}^b_3\hat{J}^c_4 + \hat{J}^a_2\hat{J}^b_3\hat{J}^c_4\big) = 0 \quad (3.34)$$

by angular momentum conservation, $\hat{J}^a_1 + \hat{J}^a_2 + \hat{J}^a_3 + \hat{J}^a_4 = 0$. Note that we can in principle 'deform' any 4-valent vertex to conform with this choice (see figure 8). Similar results hold true when yet more edges enter the vertex from one octant.

Let us now return to the coefficients $\epsilon(e_1,e_2,e_3)$ in (3.32). Though manifestly background independent they do depend on how we choose to position the faces and edges: for instance, there is no *a priori* argument that would force a given edge linking two vertices in adjacent cells to appear 'to the left' or 'to the right' of the face connecting the two vertices, as illustrated in figure 9, and in fact almost any combination of factors ± 1 is possible. To eliminate this remaining ambiguity in the definition of the volume operator, one now averages over the various possibilities [9]. More specifically, this is done by integrating over the positions of the surfaces with some measure $\mathrm{d}\mu(\theta_1,\theta_2,\theta_3)$, where the θ_i are suitable angular coordinates, viz.

$$\hat{\epsilon}(e_1,e_2,e_3) := \int \mathrm{d}\mu(\theta_1,\theta_2,\theta_3)\;\sigma(e_1,S^1_I(\theta_1))\,\sigma(e_2,S^2_I(\theta_2))\,\sigma(e_3,S^3_I(\theta_3))\,. \quad (3.35)$$

This leads to a well defined combinatorial answer depending on the relative orientation of the three edges (e_1,e_2,e_3) (= the sign of the determinant of the dreibein spanned by the tangents to the three edges at the vertex), but only modulo an overall factor which encapsulates the freedom in the choice of the measure $\mathrm{d}\mu$. In principle such a factor can arise for each vertex, but in order to preserve diffeomorphism invariance of the volume these factors should be chosen to be *the*

[9] We might note that this averaging is 'undemocratic' insofar as it is applied only to cells containing a vertex. If it were extended to other cells, some of the unwanted features which we just disposed of might re-appear.

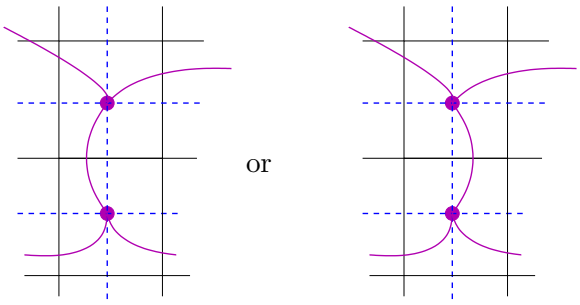

Figure 9: Sign ambiguities in the volume operator, present because the edges can be deformed and thereby moved to a different side of any of the three surfaces S_I.

same for all vertices. This factor is the one ambiguity which remains in the definition of the volume operator, and this ambiguity cannot be eliminated.

Finally, given an assignment of $\hat{\epsilon}(e_1, e_2, e_3)$ for each vertex, how do we actually compute the spectrum of the volume operator in (3.29)? First of all, it is easy to see that this operator always vanishes on three-valent vertices which satisfy the Gauss constraint. This is a direct consequence of angular momentum conservation, or equivalently

$$\epsilon_{abc}\Big(\rho_{j_{e_1}}(\tau^a)\Big)_{\alpha_1\beta_1}\Big(\rho_{j_{e_2}}(\tau^b)\Big)_{\alpha_2\beta_2}\Big(\rho_{j_{e_3}}(\tau^c)\Big)_{\alpha_3\beta_3} C^{j_{e_1}\,j_{e_2}\,j_{e_3}}_{\beta_1\beta_2\beta_3} = 0\,. \tag{3.36}$$

In other words, the volume operator vanishes on gauge invariant two- and three-valent vertices (but *not* on gauge variant ones), and picks up contributions only from spin networks with vertices of valence ≥ 4. Given any spin network wave function, the first step is to work out the action of $\hat{q}_I$ according to the above rules. In a second step this action must be diagonalised for each I on the spin network wave functions. This can be done because the operators $\hat{q}_I$ commute amongst each other, *i.e.* $[\hat{q}_I, \hat{q}_{I'}] = 0$ for $I \neq I'$. Only now can we define the absolute value $|\hat{q}_I|$ and take the square root $\sqrt{|\hat{q}_I|}$ using the spectral decomposition of this operator (note that $\hat{q}_I$ is not necessarily a positive operator). Unfortunately, these steps are all but straightforward to carry out in practice, except perhaps for very simple configurations with low-valence nodes [86, 87], and so the spectrum of the volume operator is not known exactly. It is not even known whether this spectrum is bounded away from zero (ignoring the trivial zero volume wave functions). The lack of a more explicit representation for the operators $\sqrt{\hat{q}_I}$ is also a serious impediment towards a better understanding of the Hamiltonian constraint.

3.5 Some comments on regularisation dependence

At this point some comments regarding the regularisation procedures for kinematical operators which we have outlined above, are in order. Similar comments will apply to the regularisation of the Hamiltonian constraint operator. First of all it should be evident from the foregoing discussion that the notions of 'finiteness' and 'regulator independence' as currently used in LQG are not the same as in conventional quantum field theory. There one insists that renormalised quantities should come out the same (modulo a finite number of normalisations) after removal of the regulator, *no matter how the latter is chosen*. By contrast, LQG does not claim that different prescriptions for constructing an operator will always lead to the same physical result. This is well illustrated by the volume operator, whose definition and whose very existence hinge on various choices that have been made along the way.

Let us nevertheless briefly summarise what is meant here by 'regularisation'. In order to regulate a given classical phase space quantity defined by an integral (such as the volume), one

first approximates that integral by a Riemann sum (which is of course assumed in the classical theory to converge to the desired quantity in the limit of infinite refinement). One then replaces in this finite Riemann sum all terms by operators, usually in such a way that the various area or volume elements are absorbed into the definition of these operators (as exemplified by the 'small' operators $\hat{F}^a_{S^a_I}[\tilde{E}]$ entering the definition of the area and volume operators). Acting with a sequence of these discretised operators on an arbitrary, but fixed spin network state, the resulting sequence of refinements (labelled by N, say) eventually becomes stationary (*i.e.* constant for N sufficiently large), because the state is kept fixed as the limit $N \to \infty$ is taken. This follows straightforwardly for the area operator, but requires a choice of 'preferred' regularisations for the volume operator. With these choices, the 'correct' result is always obtained already after a finite number of refinements. A good analogy would be the Riemann integral of a step function, for which a suitably refined Riemann sum attains the exact value already after finitely many steps, after which the result no longer depends on the 'regulator' (*i.e.* the size of the intervals defining the Riemann sum). In this sense, the results of LQG are indeed 'regulator independent'.

From what has just been said, it is evident that infinities can never appear in the LQG regularisation procedure, and in this sense the resulting theory is 'finite', at least as far as the kinematical operators are concerned. LQG nevertheless requires the regularisation of the area and volume operators in order even to be able to define the quantum counterparts of the classical constraints via Thiemann's trick. The outlined regularisation is, therefore, not introduced to remove divergences. Standard short distance, QFT divergences 'disappear' in the LQG approach by the very construction of the theory: all states are discrete, and at any step of the calculation one deals only with a finite number of objects. The price one pays are ambiguities of the type encountered above, some of which can only be eliminated by making *ad hoc* choices, such as the rule prescribing the way in which an edge has to traverse a cell. We shall encounter more such ambiguities when we attempt to define the Hamiltonian constraint operator.

Of course, these considerations do not necessarily imply any inconsistency. Namely, one can always adopt the above prescriptions for the kinematic operators as *definitions*, which can be plausibly related to the corresponding classical objects, and in this sense the definitions of the above operators in LQG are perfectly acceptable. What we would like to emphasise, however, is that with this construction the work is hardly begun. The far more difficult task is now to show that these definitions, when put together and inserted in the constraints, yield a sensible quantum theory, and a consistent quantum constraint algebra, in particular. The situation here is reminiscent of the one in the early days of constructive quantum field theory (see e.g. [88] and references therein). The problem then was not so much finding non-perturbative regulators and establishing the existence of selected limits when the regulator is removed. Rather, the main problem was to establish that the resulting theory has all the desired physical properties, to wit, full Poincaré invariance, locality and causality. As is well known, no interacting quantum field theory in four dimensions satisfying all these requirements is known to this day [10]. And for those low dimensional models which could actually be shown to exist and possess all the requisite properties, the tightly knit construction left no room for ambiguities other than the normalisation of a finite number of physical parameters (masses and coupling constants).

3.6 Coupling to matter fields

Any realistic theory of quantum gravity should at least allow for a coupling to the matter fields of the Standard Model of elementary particles. In the framework of geometrodynamics it is fairly easy to extend the relevant formulas to incorporate matter fields (but, of course, the difficulties

[10] Even the best candidate model for realising the quantum field theorist's dream of a UV finite theory, $N = 4$ supersymmetric Yang-Mills theory, has not been rigorously shown to exist beyond perturbation theory.

with quantisation remain the same as before). Likewise, it is possible to include gauge fields and fermions in the LQG formalism, essentially without any restriction on the gauge group or the number and structure of fermion families — as we already pointed out LQG does not appear to put any restrictions on the matter couplings (such as renormalisability). Armed with the translation rules of section 2.2, the transition from the 'old' metric or dreibein variables to the loop variables is in principle straightforward, though usually cumbersome (and, we may add, of considerably diminished esthetical appeal in comparison with Ashtekar's original reformulation of *pure* gravity). Because there is an ample literature on this subject (see e.g. [89, 90, 91, 92, 93, 94]), and because the introduction of matter couplings does not alleviate the problems with the Hamiltonian constraint in LQG, we will be very brief and only mention two selected aspects here.

Consider, for example, the coupling of fermion fields in the LQG setting [89, 95]. The starting point is the classical continuum action of a Dirac fermion coupled to gravity

$$S = \int \mathrm{d}^4x\, E\, \bar{\chi}\gamma^A E_A{}^\mu \mathcal{D}_\mu \chi \equiv \int \mathrm{d}^4x\, E\, \bar{\chi}\gamma^A E_A{}^\mu \big(\partial_\mu + \omega_{\mu BC}\gamma^{BC}\big)\chi\,, \tag{3.37}$$

where we have omitted a possible mass term for simplicity, as its inclusion would pose no new conceptual problems. There are two immediate questions one faces when trying to construct quantum Hamiltonian for this system. One is the choice of the conjugate pair of variables for fermions; finding the proper fermionic variables may also require redefinitions of the other fields so as to 'diagonalise' the canonical (Poisson or Dirac) brackets. For instance, for the above model it turns out [90] that a good choice for the conjugate pair of fermionic variables is the pair of *half-densitised* spinors

$$\tilde{\chi} \equiv \sqrt[4]{\det \tilde{E}}\,\psi\,, \quad \pi = i\tilde{\chi}^\dagger\gamma^0\,. \tag{3.38}$$

If instead of densitised spinors we used bare spinors (without any factors of the determinant of the dreibein), one would have to modify the gravitational connection [90] in order to obtain diagonal canonical brackets, but at the expense of making the connection complex again

$$\tilde{A}^a_m = -\frac{1}{2}\epsilon^{abc}\omega_{m\,bc} + \gamma K^a_m + \frac{i}{4\sqrt{\tilde{E}}} e^a_m \bar{\chi}\chi\,. \tag{3.39}$$

This leads to a new reality problem, similar to the one already encountered in [59], and not different from the one which eventually led LQG to abandon the 'old' Ashtekar variables with $\gamma = \pm i$. The second question is how to express all variables in terms of the LQG variables $h_e[A]$ and $F^a_S[\tilde{E}]$. The answer to this question can be obtained by employing the identities (2.26) and (2.27). A similar procedure will be spelled out in the next section in order to construct the Hamiltonian constraint, so we will refrain here from giving more details.

When implementing the matter Hamiltonians in terms of spin networks, the description of (abelian or non-abelian) gauge fields and matter fields is somewhat similar to the description of standard lattice gauge theories: matter fields (scalars or fermions) are attached to the vertices of the spin network, while gauge fields are associated to the edges (links), see e.g. [91, 92, 93] for a treatment of electromagnetism and [94] for scalar field theories in the LQG formalism. There are, however, two main differences between lattice gauge theory and LQG. In the former, the space or spacetime lattice is given, and hence there are only three (or four) possible directions for the edges. The coupling between two fermions χ_1 and χ_2 on adjacent vertices 1 and 2 connected by an edge $e \equiv (12)$, is uniquely given by $\bar{\chi}_1\gamma_{12}h_{12}[A]\chi_2$, where γ_{12} is the γ-matrix 'along' the edge (12). In LQG (and in general relativity), this simple formula fails because the direction of the link is arbitrary, and a vielbein factor must be inserted to 'align' the γ-matrix with the holonomy. When rewriting the dreibein by means of the formulas of section 2.2, this usually leads to rather

awkward expressions in terms of the spin network variables $h_e[A]$ and $F_S^a[\tilde{E}]$. This problem is further compounded by a second difficulty which goes back to the issue of what the correct value is for the Barbero-Immirzi parameter. As we mentioned already, for $\gamma \neq \pm i$, the connection A is not the pullback of the (chiral) spin connection. Therefore, the holonomy $h_{12}[A]$ does not effect the correct parallel transport between χ_1 and χ_2, and must accordingly be amended by further (and again rather awkward) modifications.

We should stress that, *even* when one has obtained an expression for the matter action in terms of spin network operators, one does not know whether the resulting quantum theory has any relation to the usual Fock-space quantised theory. Showing that such a relation exist requires the field-theory analogues of the 'shadow states' for point particles [21]. These states are complicated linear combinations of elementary spin network states, approximating Fock space coherent states. For fermionic theories, these are not known (some results have been obtained in the case of Maxwell [91, 92, 93] and scalar fields [94]). When gravity is included, the situation becomes even more complicated [22, 23, 24, 25, 26, 27].

A completely different question concerns the role of the structure group SU(2), and the possibility of generalising Ashtekar's variables to higher dimensions, to achieve a Kaluza-Klein type unification of matter and gravity. The closure of the classical constraint algebra in the Ashtekar reformulation of gravity requires use of the identity $\epsilon_{abe}\epsilon^{cde} = 2\delta^{cd}_{ab}$, which is only valid for the structure constants of the group SU(2) and its non-compact form SO(1,2). These properties no longer hold for SU(N) when $N > 2$. For this reason, the LQG reformulation of gravity works only in three and four spacetime dimensions. Nevertheless, one can try to generalise the formalism to SU(N) and thereby arrive at a novel type of coupling between gravity and Yang-Mills fields [96, 97]. A very different attempt to generalise the formalism to include matter couplings is based on the observation that an inverse densitised vielbein also appears naturally in certain reformulations of $D = 11$ supergravity [98]. It is not known whether any of these *ansätze* can be implemented in a spin network formulation (although replacing SU(2) holonomies by SU(N) holonomies seems straightforward enough).

4 Quantum constraints

As we have seen in section 2.1, the Hamiltonian framework leads to three classical constraints: the Gauss, diffeomorphism and Hamiltonian constraints. To implement these constraints at the quantum level, one must first properly *define* them, *i.e.* express them in terms of the elementary variables, the holonomies and fluxes, and then investigate their properties. We will first discuss the kinematical constraints, which are 'easy' (relatively speaking), whence we can be brief. We will then turn to the truly difficult part of the story – the Hamiltonian constraint.

4.1 Kinematical constraints

The Gauss constraint is already expressed in canonical variables $h_e[A]$ and $F_S[\tilde{E}, f]$, and one can solve it exactly. This is simply because fulfilment of the Gauss constraint is equivalent to the gauge invariance of the spin network wave function: invariance is simply achieved by contracting the SU(2) representation indices entering a given vertex in an SU(2) invariant manner. Equivalently, one says that *angular momentum is conserved at each vertex*, viz.

$$\sum_e \hat{J}^a_{(e,v)} = 0 \tag{4.1}$$

for each $v \in \Gamma$. Recall (see the discussion around (3.17)) that the operator $\hat{J}_a^{(e,v)}$ is *defined* to act by inserting a matrix $\rho_{j_e}(\tau^a)$ at the vertex v in the representation belonging to the (oriented)

edge e, with appropriately contracted SU(2) representation indices, and appropriate orientation. For the following discussion, we will tacitly assume that the wave function satisfies the Gauss constraint, whenever appropriate.

The diffeomorphism constraint is more difficult to impose — unlike in geometrodynamics, one cannot immediately write down formal expressions which are manifestly diffeomorphism invariant, because the spin network functions are not supported on all of Σ, but only on one-dimensional links. Nevertheless, it would seem most natural at first sight to try to solve the diffeomorphism constraint on $\mathcal{S}$ or the kinematical Hilbert space $\mathcal{H}_{\rm kin}$. However, this is impossible for two reasons. First of all, the diffeomorphism generator does not even exist as an operator: there is no way to differentiate finite diffeomorphisms so as to obtain 'infinitesimal' ones due to the lack of weak continuity (and *a fortiori*, differentiability). Secondly, with the exception of the trivial state $\Psi = \mathbf{1}$ (the empty spin network), there is just *no* diffeomorphism invariant state in $\mathcal{H}_{\rm kin}$. This is intuitively clear because every spin network state changes under a diffeomorphism which moves the associated graph around.[11] For these reasons, LQG invokes a variant of the so-called 'group averaging' method, but in a way that is different from the standard group averaging procedure. Namely, application of the latter would require knowing an integration measure $\mathrm{d}\mu(\phi)$ on the infinite dimensional group Diff(Σ). The formally diffeomorphism invariant state would then be

$$\Psi^{\rm inv}_{\Gamma,\psi}[A] = \int_{\rm Diff(\Sigma)} \mathrm{d}\mu(\phi)\, \Psi_{\Gamma,\psi}[A \circ \phi]\,, \tag{4.2}$$

and if the measure existed, this state would legitimately belong to $\mathcal{H}_{\rm kin}$. Unfortunately, for the infinite dimensional group Diff(Σ) no such measure is known (and probably none exists), hence the above expression is devoid of meaning — confirming the conclusion above that $\mathcal{H}_{\rm kin}$ does not admit non-trivial solutions. LQG circumvents this difficulty by implementing 'group averaging' in a way that is more adapted to the non-separable Hilbert space and the properties of the scalar product (3.7). The key observation here is that the constraint can be formally solved on a much larger space, namely *the dual* $\mathcal{D}^*$ *of some dense subspace* $\mathcal{D}$ of $\mathcal{H}_{\rm kin}$, such that

$$\mathcal{D} \subset \mathcal{H}_{\rm kin} \subset \mathcal{D}^* \tag{4.3}$$

(we cannot take $\mathcal{D} = \mathcal{H}_{\rm kin}$, because then the dual space would be the same). Typically, the space $\mathcal{D}^*$ is thus a space of distributions. This is not as outlandish as it may sound at first, because for constrained systems, the solutions of the constraints are generically distributions, see e.g. [99] [12]. Given any operator $\mathcal{O}$ on $\mathcal{D}$, one must therefore also consider the adjoint operator $\mathcal{O}^*$, whose action on a particular element $\mathcal{X} \in \mathcal{D}^*$ is defined by requiring

$$\langle \mathcal{O}^*\mathcal{X} \,|\, \Psi\rangle = \langle \mathcal{X} \,|\, \mathcal{O}\Psi\rangle \tag{4.4}$$

to hold for all $\Psi \in \mathcal{D}$. This applies in particular to the Hamiltonian constraint operator. If the operator and its adjoint no longer act in the same space, the notion of *hermiticity* no longer makes sense. This may entail other difficulties of a more conceptual nature, even if one restricts attention to the 'zero eigenvalue subspace' of the (adjoint of the) Hamiltonian in $\mathcal{D}^*$ only.

The obvious ('canonical') choice for $\mathcal{D}$ is

$$\mathcal{D} = \mathcal{S} \tag{4.5}$$

[11] This problem disappears altogether in a *topological theory*, such as (2+1)-dimensional gravity, where the connection is flat and diffeomorphism invariance is manifest from the outset.

[12] A simple example is the solution of the Hamiltonian constraint for a relativistic point particle – the Klein Gordon equation – which is $\phi(p) = \delta(p^2 - m^2)$. As is well known, a scalar product and a Hilbert space can be defined on this constraint hypersurface, but the scalar product is not positive definite. These features can also be made completely explicit in more complicated, but exactly solvable theories, such as pure supergravity in three dimensions in the connection formalism, see [100].

that is, the space of finite linear combinations of spin network states. However, there may exist other viable choices for $\mathcal{D}$. The dual space $\mathcal{S}^*$ is much larger than $\mathcal{H}_{\rm kin}$, and actually too large for our purposes. We are here interested only in a certain subspace of $\mathcal{S}^*$ and its completion w.r.t. to the norm (4.8) below. This subspace is generated by averaging spin network states. More specifically, for any $\Psi_\Gamma \in \mathcal{S}$ associated with a particular graph Γ we define the 'group average' by the formal sum

$$\eta(\Psi_\Gamma) \equiv \overline{\Psi}_\Gamma := \sum_{\phi \in {\rm Diff}(\Sigma|\Gamma)} \phi^* \circ \Psi_\Gamma \tag{4.6}$$

where $(\phi^* \circ \Psi)[A] := \Psi[A \circ \phi]$, and where ${\rm Diff}(\Sigma|\Gamma)$ consists of the subgroup of ${\rm Diff}(\Sigma)$ obtained by dividing out the diffeomorphisms leaving invariant the graph Γ on which Ψ_Γ lives. While $\Psi_\Gamma \in \mathcal{S}$ by assumption, averaging the state in this way throws it out of $\mathcal{S}$ and also $\mathcal{H}_{\rm kin}$. At first sight, this seems to make matters worse: we have replaced the ill-defined integral over ${\rm Diff}(\Sigma)$ by a highly formal continuous sum. However, the day (and LQG) is now saved by two facts: (i) unlike the integral, the sum no longer requires unavailable detailed knowledge of the properties of ${\rm Diff}(\Sigma)$, and (ii) when applying an averaged spin network state $\eta(\Psi_{\Gamma'})$ (considered as a bra-vector) to any other spin network state $\Psi_\Gamma[A]$, the sum

$$\langle \eta(\Psi_{\Gamma'}) \,|\, \Psi_\Gamma \rangle = \sum_{\phi \in {\rm Diff}(\Sigma|\Gamma')} \langle \phi^* \circ \Psi_{\Gamma'} \,|\, \Psi_\Gamma \rangle \tag{4.7}$$

consists at most of a *finite number* of terms, and it is this fact which ensures that $\eta(\Psi_\Gamma)$ is indeed well defined as an element of the dual $\mathcal{S}^*$. Namely, there is no contribution at all if Γ and Γ' are not diffeomorphic (that is, cannot be made to coincide by a diffeomorphism). For Γ and Γ' diffeomorphic, only that term contributes for which Γ and $\phi^* \circ \Gamma^*$ coincide. In the above equation, the bracket on the r.h.s. is just the scalar product (3.7), while the bracket on the l.h.s. denotes the dual pairing between $\mathcal{S}$ and $\mathcal{S}^*$.

We can now promote the space of averaged diffeomorphism invariant spin network states to a pre-Hilbert space by 'dividing out the volume of the gauge group' (= ${\rm Diff}(\Sigma)$). More precisely, we define the scalar product between two such states as

$$\langle\!\langle \eta(\Psi_{\Gamma',\psi'}) \,\big|\, \eta(\Psi_{\Gamma,\psi}) \rangle\!\rangle := \langle \eta(\Psi_{\Gamma',\psi}) \,\big|\, \Psi_{\Gamma,\psi} \rangle \,. \tag{4.8}$$

Here the bracket on the r.h.s. is the dual pairing between $\mathcal{S}$ and $\mathcal{S}^*$, and the bracket $\langle\!\langle .|. \rangle\!\rangle$ on the l.h.s. denotes the new scalar product. The new Hilbert space $\mathcal{H}_{\rm diff}$ of diffeomorphism averaged states is then obtained by completion w.r.t. this norm, and can be thought of to consist of the averaged versions of the states (3.8). Observe that the Hilbert space $\mathcal{H}_{\rm diff}$ is still a 'small' subspace of $\mathcal{S}^*$ because the vast majority of elements in $\mathcal{S}^*$ does not correspond to diffeomorphism averaged spin network states, and cannot be obtained as limits of such states w.r.t. to the above norm. Moreover, $\mathcal{H}_{\rm diff}$ is 'smaller' than the original kinematical Hilbert space $\mathcal{H}_{\rm kin}$, because distinct elements of the latter which are related by a diffeomorphism get mapped to the *same* element of $\mathcal{S}^*$; equivalently, the averaging map has a very large kernel (while $\mathcal{H}_{\rm kin}$ is non-separable, it is expected that $\mathcal{H}_{\rm diff}$ *is* separable [101, 102] provided one averages over diffeomorphisms which are not necessarily analytic). There is no way to extend the above norm from $\mathcal{H}_{\rm diff}$ to all of $\mathcal{S}^*$, in accordance with the fact that distribution spaces cannot be made into Hilbert spaces, but at best into topological vector spaces (but there is no consensus as to which of the many possible topologies is the right one for $\mathcal{S}^*$). The art here – as with any other constrained quantum mechanical system – is to find a suitable subspace of physically relevant elements, on which a scalar product can be defined.

Which, then, is the arena, or 'habitat', in which quantum gravity takes place according to LQG, and where one must ultimately consider the action of the Hamiltonian constraint?

The foregoing discussion shows that there is no easy answer to this question, and apparently also no complete consensus in the community as to which space is 'right'. The obvious first choice, the kinematical Hilbert space $\mathcal{H}_{\rm kin}$, is ruled out from the start because it does not admit diffeomorphism invariant states. The Hilbert space $\mathcal{H}_{\rm diff}$ of diffeomorphism averaged states is certainly a better option for implementing and solving the Hamiltonian constraint [29], but it, too, may not be good enough. The reason is that the action of the Hamiltonian constraint in general does not preserve $\mathcal{H}_{\rm diff}$; in other words, it maps diffeomorphism invariant states into ones that are not invariant. Consequently, when checking the commutator of two Hamiltonian constraints, one must again pass to a larger space, which is preserved by the action of the Hamiltonian, but which is no longer a Hilbert space (an explicit example are the 'vertex-smooth' states introduced in [30]). We refrain here from introducing yet more spaces, and yet more notation (and thereby confuse readers further). We will return to this issue in the following chapter, but let us first try ourselves at the more mundane task of working out explicit expressions for the Hamiltonian.

4.2 Hamiltonian constraint

As we explained, the diffeomorphism constraint is imposed not in terms of an operator constraint, but through the formal group averaging method. By contrast, the Hamiltonian constraint is imposed as an *operator* constraint on spin network states. Hence, as a zeroth step, one has to construct this Hamiltonian operator. As we explain in more details below, the strict derivation of the Hamiltonian operator from the classical expression is lacking. What one does in practice, is to *postulate* the quantum Hamiltonian operator through its action on a generic spin network state. This *definition* is indeed motivated by the classical expression for the Hamiltonian (expressed in loop variables), but it is plagued by an infinite number of ambiguities. To test the proposed prescription and compare it with other possibilities, one would need to find states which are annihilated by the Hamiltonian and check that in the semi-classical limit these states produce sensible result. This is, however, currently out of reach: due to the great complexity of the proposed Hamiltonian, not a single physically interpretable eigenstate is known. While this should not come as a big surprise (after all, we have no reason to expect to be able to solve the theory exactly), it is remarkable that even the far more modest goal of working out in complete detail the action of the Hamiltonian on a general spin network wave function appears to be out of reach.

Even if it were possible to work out the action of the Hamiltonian on any given element of $\mathcal{S}$, we would not be done. As we will explain below, we will be mostly dealing with a *regulated Hamiltonian* $\hat{H}[N,\epsilon]$. Here, the regularisation parameter $\epsilon > 0$ enters via a plaquette $P(\epsilon)$, which is attached to a vertex and which must be shrunk to zero at the end of the calculation. However, the limit $\epsilon \to 0$ of $\hat{H}[N,\epsilon]$ does *not* exist on $\mathcal{S}$, because wave functionals supported on the same network, but with an extra loop $\partial P(\epsilon)$ attached to one of the vertices are orthogonal to one another for different values of ϵ by (3.7). For this reason, and for the reasons explained in the preceding subsection, we must transfer the action of the Hamiltonian to the dual space $\mathcal{S}^*$, or at least to some physically motivated subspace thereof, and define the limit $\epsilon \to 0$ for the adjoint operator by means of (4.4), that is, by demanding

$$\langle \hat{H}^*[N]\mathcal{X} \,|\, \Psi\rangle = \lim_{\epsilon\to 0} \langle \mathcal{X} \,|\, \hat{H}[N,\epsilon]\Psi\rangle \tag{4.9}$$

to hold for all $\Psi \in \mathcal{S}$. This means that the dual Hamiltonian $\hat{H}^*[N]$ can only be defined as a *weak limit*, which moreover requires simultaneous knowledge of the action of the original regulated Hamiltonian $\hat{H}[N,\epsilon]$ on *all* elements of $\mathcal{S}$. The hope is that by transferring the action of $\hat{H}$ to the space of diffeomorphism invariant states, and by making the dual operator 'diffeomorphism covariant' as in [30], a well behaved limit may be obtained. Unfortunately, *explicit* expressions

for this operator that could be used for any practical application of the formalism are even more difficult to obtain than for the Hamiltonian acting on spin network states, and certainly beyond the scope of our modest efforts in this section. Nevertheless, we should stress once more that the dual Hamiltonian $\hat{H}^*$ is our true object of interest, and that all further considerations (constraint algebra, semi-classical states, and so on) should be based on it, rather than on the regulated Hamiltonian acting on $\mathcal{S}$. With these caveats in mind, let us now proceed.

To motivate the form of the *quantum Hamiltonian* one starts with the classical expression, written in loop variables. Despite the simplifications brought about by equation (2.21), the Hamiltonian constraint still looks formidable,

$$H[N] = \int_\Sigma \mathrm{d}^3x\, N \frac{\tilde{E}^m_a \tilde{E}^n_b}{\sqrt{\det \tilde{E}}} \Big(\epsilon^{abc} F_{mnc} - \frac{1}{2}(1+\gamma^2) K_{[m}{}^a K_{n]}{}^b \Big) \,. \tag{4.10}$$

In order to write the constraint in terms of only holonomies and fluxes, one has to eliminate the inverse square root as well as the extrinsic curvature factors. This can be done using the relations (2.26)–(2.28). Inserting these into the Hamiltonian constraint one obtains the expression

$$H[N] = \int_\Sigma \mathrm{d}^3x\, N \epsilon^{mnp} \,\mathrm{Tr}\, \Big(F_{mn} \{A_p, V\} - \frac{1}{2}(1+\gamma^2) \{A_m, \bar{K}\} \{A_n, \bar{K}\} \{A_p, V\} \Big) \,. \tag{4.11}$$

This expression is the starting point for the construction of the quantum constraint operator.

The first remark which should be made concerns the presence of the volume operator in (4.11). Although the square V^2 has been expressed in terms of fluxes in section 3.4.2, the spectrum of this operator is not known. Therefore, it is in practice not possible to take the square root, and as a consequence it is not known how V acts on a spin network. In any analysis of the Hamiltonian constraint, it is therefore only possible to use certain qualitative properties of the volume operator. For quantitative results, it would be necessary to first diagonalise V^2.

The next step is to express the Ashtekar connection A_m as well as field strength F_{mn} and the extrinsic curvature $\bar{K}$ in terms of holonomies. As in lattice gauge theory, this requires the introduction of a cell structure. In the *classical* expression one splits the target space into cells of size $\sim \epsilon^3$ in *coordinate* space. In order to express connections and curvatures in terms of holonomies, one now has to choose edges e_m within each cell, along which the holonomies are going to be evaluated. Once the edges e_m are chosen, one uses that the connection along a line of *coordinate* distance ϵ can be approximated by

$$\begin{aligned} \Big\{ \int_e A, V \Big\} &= -h_e[A]^{-1} \big\{ h_e[A], V \big\} + o(\epsilon) \,, \\ \Big\{ \int_e A, \bar{K} \Big\} &= -h_e[A]^{-1} \big\{ h_e[A], \bar{K} \big\} + o(\epsilon) \,. \end{aligned} \tag{4.12}$$

For the field strength one has the standard lattice expression

$$\int_{P(\epsilon)} F[A] = \frac{1}{2} \Big(h^{-1}_{\partial P(\epsilon)}[A] - h_{\partial P(\epsilon)}[A] \Big) + o(\epsilon^3) \,, \tag{4.13}$$

where $P(\epsilon)$ is a two-dimensional infinitesimal plaquette with boundary $\partial P(\epsilon)$, that shrinks to zero as $\epsilon \to 0$. The extrinsic curvature is expressed using the relation (2.28), in which in one substitutes (4.13). Finally, as for the volume operator, one replaces integrals with Riemann sums over the cells, which we will again label with α. After replacing Poisson brackets by commutators, the resulting *regulated Hamiltonian* reduces to a sum over the vertices v_α of the spin network

with lapses $N(v_\alpha)$

$$\hat{H}[N,\epsilon] = \sum_\alpha N(v_\alpha)\,\epsilon^{mnp}\,\mathrm{Tr}\left(\left(h_{\partial P_{mn}(\epsilon)} - h^{-1}_{\partial P_{mn}(\epsilon)}\right) h_p^{-1}\left[h_p, V\right]\right)$$
$$+ \frac{1}{2}(1+\gamma^2)\sum_\alpha N(v_\alpha)\,\epsilon^{mnp}\,\mathrm{Tr}\left(h_m^{-1}\left[h_m,\bar{K}\right] h_n^{-1}\left[h_n,\bar{K}\right] h_p^{-1}\left[h_p,V\right]\right). \quad (4.14)$$

If one does not want to make reference to any particular spin network, the Hamiltonian can also be expressed more abstractly as a continuous sum

$$\hat{H}[N,\epsilon] = \sum_{\mathbf{x}\in\Sigma} N(\mathbf{x})\hat{H}(\mathbf{x};\epsilon) \quad (4.15)$$

which, on any given spin network, reduces to a finite sum as in (4.14). In the remainder, we will keep the regularisation parameter $\epsilon > 0$ fixed, and not indicate it explicitly any more in the formulas below.

The classical expression (4.10) is independent of the coordinate system chosen on Σ. Once we pass to the quantum theory, a generic spin network state on which the quantum Hamiltonian acts is *not* diffeomorphism invariant. Hence, in contrast, the action of the quantum expression (4.14) on a generic spin network state *does* depend strongly on the choice of the coordinate system on Σ. To be more precise, the action depends on whether or not the basis vectors (*i.e.* the directions m,n and p) are aligned with the spin network edges at a given node. There exists a large number of possibilities for these alignments, and one might simply define the action of the Hamiltonian to employ a particular alignment. Some of these possibilities are explicitly listed below. A priori, it is not clear (to us) whether these different choices correspond to different "gauge fixing" choices, or whether they are genuine ambiguities in the quantisation procedure. If it is the first case, then one should show that the final result, namely the observables quantities and the form of the algebra, are *independent* of the way in which one makes the choice. In the latter case, the consistency requirements and the comparison with experiments should discriminate between various options (see the very end of this section).[13]

To get an idea of the complications, let us consider the action of the first term in the regulated Hamiltonian constraint on a spin network, for given $\epsilon > 0$. Some of the possibilities in which this term may act are illustrated in figure 10. Here we have chosen a particular coordinate system on Σ, and focus on a term in the sum over m,n and p for which the directions m and n are aligned with two of the edges, while the direction p is orthogonal but not aligned with any existing edge. The action of $\hat{H}$ now consists of three steps.

In the first step, the operator $h_p^{-1}[h_p, V] = V - h_p^{-1}Vh_p$ acts at the vertex of the network. When V in the second term acts on three edges already present in the spin network, the result cancels against the first term. However, when V acts on two existing edges and on the h_p factor, the result is non-trivial. The vertex gets 'opened', and the two oppositely oriented edges h_p and h_p^{-1} edges are glued in. The su(2) representation j of these edges is not a priori fixed, and represents an ambiguity in the definition of the Hamiltonian. The addition of these new edges changes the spin network wave function at the edge to

$$C^{j_1 j_2 j_3}_{\alpha_1\alpha_2\alpha_3}\left(h^{j_1}\right)_{\alpha_1\beta_1}\left(h^{j_2}\right)_{\alpha_2\beta_3}\left(h^{j_3}\right)_{\alpha_3\beta_3}$$
$$\to \left(h_p^{-1}\right)_{\lambda_2\lambda_3}\left(\tilde{C}^{j_2 j_3 j}_{\alpha_2\alpha_3\alpha_1}\left(h^{j_2}_m\right)_{\alpha_2\beta_2}\left(h^{j_3}_n\right)_{\alpha_3\beta_3}\left(h^{j}_p\right)_{\alpha_1\lambda_1}\right)\left(h^{j_1}_q\right)_{\lambda_3\beta_1} + 2 \text{ terms}. \quad (4.16)$$

[13] Compare this to the case of the volume operator, where diffeomorphism averaging was invoked at the level of the operator. In principle one could say that one should not need to do this averaging at all.

Here $\tilde{C}$ is a new intertwiner, obtained by the action of the volume operator on two of the original edges emanating from v_α as well as on the holonomy h_p. As we will see, one has to distinguish between the situation in which h_p is connected to the m or n edge (figure 10a, corresponding to the 2nd and 3rd term above) and the situation in which this is not the case (figure 10b, corresponding to the first term above).

In the second step, the plaquette $(F_{mn})_{\lambda_1\lambda_2}$ is connected to the spin network. The resulting networks are again displayed in figure 10. When the lengths of the edges h_p and h_p^{-1} are taken to zero, the plaquette lies in the plane spanned by two of the edges. The difference between the two cases becomes more visible in the third step. There, one uses the fact that the insertion of the plaquette has led to double edges, *i.e.* products of holonomies along the same edge. These products can be reduced by using the tensor-product decomposition

$$\big(h_e^{j_1}\big)_{\alpha_1\beta_1}\big(h_e^{j_2}\big)_{\alpha_2\beta_2} = \sum_k C^{j_1 j_2 k}_{\alpha_1\alpha_2\gamma_1} C^{j_1 j_2 k}_{\beta_1\beta_2\gamma_2} \big(h_e^k\big)_{\gamma_1\gamma_2} . \tag{4.17}$$

Applying this expression twice to the situation in figure 10b, the spin network becomes

$$\begin{aligned}
&\tilde{C}^{j_2 j_3 j}_{\alpha_2\alpha_3\alpha_1}\bigg(\big(h_m^j\big)_{\alpha_1\xi}\big(h_m^{j_2}\big)_{\alpha_2\sigma_1}\bigg)\big(h_m^j\big)_{\xi\rho}\bigg(\big(h_n^j\big)_{\lambda_3\rho}\big(h_n^{j_3}\big)_{\alpha_3\sigma_2}\bigg)\big(h_m^{j_2}\big)_{\sigma_1\beta_2}\big(h_n^{j_3}\big)_{\sigma_2\beta_3}\big(h^{j_1}\big)_{\lambda_3\beta_1} \\
&\quad = \tilde{C}^{j_3 j_2 j}_{\alpha_2\alpha_3\alpha_1}\bigg(\sum_{k_1} C^{j j_2 k_1}_{\alpha_1\alpha_2\rho_1}\big(h^{k_1}\big)_{\rho_1\rho_2} C^{j j_2 k_1}_{\xi\sigma_1\rho_2}\bigg)\big(h^j_{mn}\big)_{\xi\phi}\bigg(\sum_{k_2} C^{j j_3 k_2}_{\lambda_3\alpha_3\eta_1}\big(h^{k_2}\big)_{\eta_1\eta_2} C^{j j_3 k_2}_{\phi\sigma_1\eta_2}\bigg) \\
&\quad = \sum_{k_1,k_2}\bigg(V^A_{\lambda_3\rho_1\eta_1}\big(h_q^{j_1}\big)_{\beta_1\lambda_3}\big(h_m^{k_1}\big)_{\rho_1\rho_2}\big(h_n^{k_2}\big)_{\eta_1\eta_2}\bigg) V^B_{\xi\sigma_1\rho_2}\big(h^j_{mn}\big)_{\xi\phi}\big(h_m^{j_2}\big)_{\rho_2\beta_2}\; V^C_{\phi\sigma_1\eta_2}\big(h_m^{j_3}\big)_{\eta_1\beta_3} .
\end{aligned} \tag{4.18}$$

The Clebsch-Gordan decomposition has modified the original vertex to a new vertex V^A, and has also introduced two new vertices V^B and V^C,

$$V^A_{\lambda_3\rho_1\eta_1} = \tilde{C}_{\alpha_2\alpha_3\alpha_1} C_{\alpha_1\alpha_2\rho_1} C_{\lambda_3\alpha_3\eta_1} , \qquad V^B_{\xi\sigma_1\rho_2} = C_{\xi\sigma_1\rho_2} , \qquad V^C_{\phi\sigma_2\eta_2} = C_{\phi\sigma_2\eta_2} . \tag{4.19}$$

Thus the net result of this insertion of the plaquette amounts to adding a new edge in an ϵ-neighbourhood of the initial vertex. A similar analysis can be performed to deduce the action in case of figure 10a.

Let us now comment on other possible coordinate systems and alignments. The use of the tensor product decomposition (4.17) depends crucially on the edges of the spin network being aligned with the edges of the plaquette (and, through (4.14), with the coordinate system on Σ). If this is not the case, the plaquette will not get glued into the spin network, but instead remains freely floating, only connected at the original vertex (see figure 11). In the LQG literature, such cases are excluded by hand, by always choosing a local coordinate system around each vertex which is completely aligned with (three of the) edges emanating from the vertex. The resulting, *postulated* quantum Hamiltonian becomes

$$\begin{aligned}
\hat{H} = &\sum_\alpha \sum_{I,J,K} N(v_\alpha)\, L^{IJK}\, \mathrm{Tr}\Big(\big(h_{\partial P_{IJ}} - h^{-1}_{\partial P_{IJ}}\big)\, h^{-1}_{s_K}\big[h_{s_K}, V\big]\Big) \\
&+ \frac{1}{2}(1+\gamma^2)\sum_\alpha \sum_{I,J,K} N(v_\alpha)\, L^{IJK}\, \mathrm{Tr}\Big(h^{-1}_{s_I}\big[h_{s_I}, \bar{K}\big]\, h^{-1}_{s_J}\big[h_{s_J}, \bar{K}\big]\, h^{-1}_{s_K}\big[h_{s_K}, V\big]\Big) .
\end{aligned} \tag{4.20}$$

The second sum is a sum over all triplets of edges emanating from the vertex v_α in the centre of the cell α. The boundary of a plaquette in the plane spanned by the edges s_I and s_J is denoted by ∂P_{IJ}. Finally, the value of the constant, fully anti-symmetric tensor L^{IJK} depends on the particular vertex α.

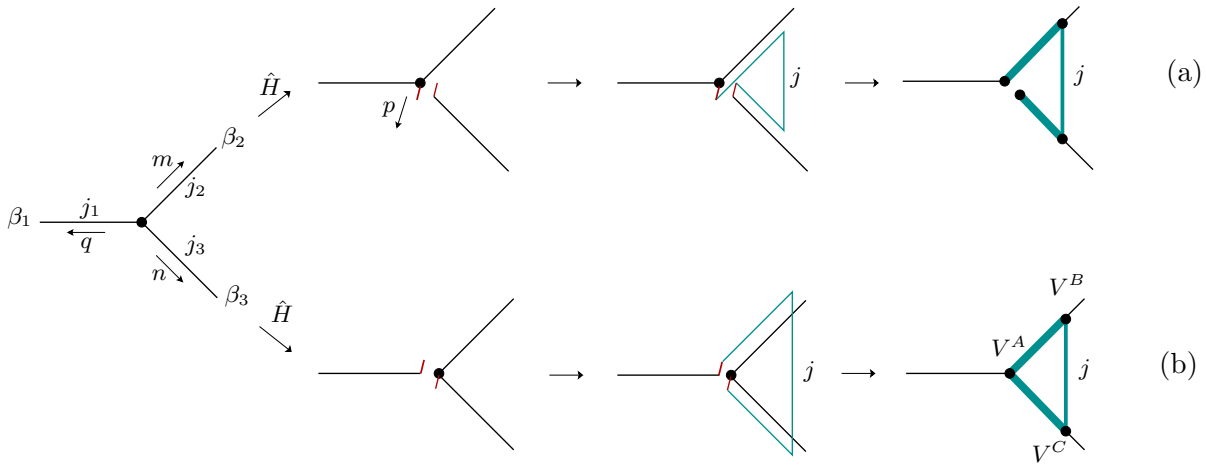

Figure 10: The action of the first term in the Hamiltonian constraint on a spin network, in the special case that the m and n directions are aligned with two edges emanating from the node. The two small edges in the p direction, introduced in the first step, stick out of the plane in which the plaquette extends.

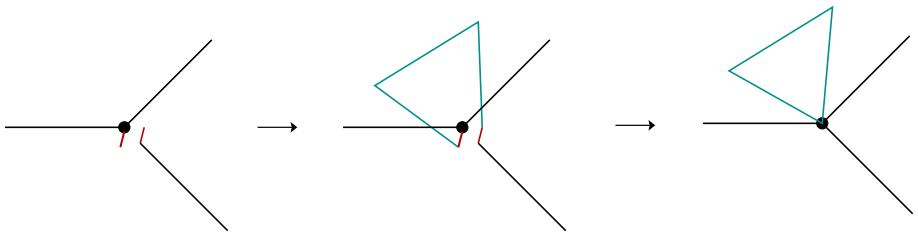

Figure 11: The action of the first term in the Hamiltonian constraint, for the case in which none of the coordinates on Σ is aligned with any of the edges emanating from the vertex.

We have so far not discussed the second term in the constraint (4.14), which arose from the product of extrinsic curvatures in (4.10). This term is far more complicated than the first one. Its complexity stems from the complexity of the intermediate expression

$$[A_I, \bar{K}] = h_{e_I}^{-1} \left[h_{e_I}, \epsilon^{JKL} \left[(h_{\alpha_{JK}} - h_{\alpha_{JK}}^{-1}) h_{e_L}^{-1} \left[h_{e_L}, V\right],\ V\right]\right] , \tag{4.21}$$

which is obtained using (2.26)–(2.28). One recognises the combination of operators which opens up the vertex and inserts a plaquette, but there there is an additional volume operator which applies further changes to the node. If the coordinate system on Σ is aligned in such a way that all three directions coincide with an existing edge, the action of the second term in the constraint is roughly to glue in three new edges, introducing six new vertices. But as the discussion above shows, there is a plethora of other possibilities.

Let us, for clarity, summarise the problems which we see with the Hamiltonian constraint.

- *Ambiguities.* The construction of the Hamiltonian outlined here is plagued with a variety of ambiguities. Firstly, a choice has to be made for the coordinate system on Σ. The action of the Hamiltonian depends in a qualitative way on this choice. Secondly, there is an ambiguity in choosing the representation in which the trace is evaluated (*i.e.* the representation in which the field strength F_{mn} and the holonomy h_p are taken). Thirdly, there are operator ordering ambiguities. Finally, all ambiguities present in the construction of the volume operator are inherited by the Hamiltonian constraint.

- *Ad-hoc choices.* The quantum constraint should be well-defined and unambiguous on the full space of spin-network states. If this is not the case, it should be shown that the choices do not matter for the physical predictions, or alternatively that a particular choice is singled out by physical consistency. By focussing only on the operator (4.20), it becomes impossible to tell whether the constraint operator is indeed independent of the coordinate choice on Σ.

- *Ultralocality.* From the explicit discussion of the Hamiltonian constraint it should be clear that its action is always 'ultralocal', in the sense that all changes to the spin network are made in an $\epsilon \to 0$ neighbourhood of a given vertex, while the spin network graph is kept fixed [103, 104, 105]. More specifically, it has been argued [30] that the Hamiltonian acts at a particular vertex only by changing the intertwiners at that vertex, viz.

$$\hat{H}(v)\,|\Gamma, \{j\}, \{C\}\rangle = \sum_{I,J} U^a |\Gamma, \{j\}, C_a(\Gamma, \{j\}, I, J)\rangle \tag{4.22}$$

  where U^a is the traceless part of the holonomy associated with the plaquette P_{IJ} and C_a a triplet of new intertwiners at v, and the sum is over pairs of edges I, J entering the vertex. This is in contrast to what happens in lattice field theories. There the action of the Hamiltonian always non-trivially links two existing nodes, the plaquettes are by construction always spanned between existing nodes, and the continuum limit involves the lattice as a whole, not only certain sub-plaquettes. It is by no means clear whether diffeomorphism averaging can rid us of this problem.

5 Constraint algebra

As we already emphasised in the introduction, the proper closure of the quantum constraint algebra is perhaps *the* crucial consistency requirement that must be met by any candidate theory of quantum gravity. Of course, the devil hides in the meaning of the word 'proper'. To find out where it might possibly hide, we now return to the points raised at the end of section 4.1. We will not bother here with the commutators involving the kinematical constraints which are straightforward to take care of in any approach. The crucial relation is the commutator of two Hamiltonian constraints, whose classical analog reads

$$\big\{H_0[M]\,,\, H_0[N]\big\} = \int_\Sigma \mathrm{d}^3x\,(M\partial_m N - N\partial_m M) g^{mn} H_m\,. \tag{5.1}$$

Here $H_m \equiv e_m{}^a H_a$ is the diffeomorphism generator and the smeared Hamiltonian is defined as

$$H_0[M] := \int_\Sigma \mathrm{d}^3x\, H_0(\mathbf{x}) M(\mathbf{x})\,. \tag{5.2}$$

To be sure, the proper closure of the quantum constraint algebra, *i.e.* the issue of promoting (5.1) to a relation between operators acting on a suitable Hilbert space of states, is also one of the major unsolved problems of the conventional geometrodynamics approach. There it has not been possible to resolve the ordering and other ambiguities, and to get rid of the infinities resulting from the clash of functional differential operators (the situation is partly better, though, in certain supergravity models whose constraint superalgebra always contains the bosonic constraint algebra as a subalgebra [59]). Here we simply wish to explain why the ansätze available in the LQG literature so far seem hardly more satisfactory than what conventional geometrodynamics has to say on this issue.

5.1 The bosonic string as an example

Before continuing our discussion of the quantum constraint algebra in canonical gravity, we would like to recall that there does exist a model which can be solved exactly, and all the way. Hence it can serve as an example of how quantisation should work. This model is the bosonic string (in a Polyakov type formulation), when viewed as a *bona fide* model of matter coupled quantum gravity in two dimensions (see also [106] for an earlier discussion of the bosonic string in this context). For this purpose, we only need to forget about the target space interpretation and its ramifications, such as string scattering, or higher order corrections in g_s, involving splitting and joining of strings (in the present context, the latter processes would amount to changes of the spatial topology, 'baby universes' and the like, hence, 'third quantisation'). The resulting theory is manifestly independent of the two-dimensional space-time background, free of divergences and other pathologies, and also possesses nice observables [14] (vertex operators). More specifically, the diffeomorphism and scalar constraints D and H are given by (see e.g. [107] or [108])

$$\begin{aligned} D[f] &= \frac{1}{2}\int_0^{2\pi} \mathrm{d}\sigma\, f(\sigma) P_\mu(\sigma)\partial_\sigma X^\mu(\sigma)\,, \\ H[N] &= \frac{1}{4}\int_0^{2\pi} \mathrm{d}\sigma\, N(\sigma)\Big(P_\mu(\sigma)P^\mu(\sigma) + \partial_\sigma X_\mu(\sigma)\partial_\sigma X^\mu(\sigma)\Big)\,. \end{aligned} \tag{5.3}$$

with shift $f(\sigma)$ and lapse $N(\sigma)$. They give rise to the classical constraint algebra

$$\begin{aligned} \Big\{D[f_1],\ D[f_2]\Big\}_{\mathrm{PB}} &= \tfrac{1}{2}D[f_1\partial f_2 - f_2\partial f_1]\,, \\ \Big\{D[f],\ H[N]\Big\}_{\mathrm{PB}} &= \tfrac{1}{2}H[f\partial N - N\partial f]\,, \\ \Big\{H[N_1],\ H[N_2]\Big\}_{\mathrm{PB}} &= \tfrac{1}{2}D[N_1\partial N_2 - N_2\partial N_1]\,. \end{aligned} \tag{5.4}$$

Canonical quantisation is straightforward in the conformal gauge. The state space of the theory $\mathcal{F}$ is simply the tensor product of the space of one-particle wave functions and an infinite number of harmonic oscillator Hilbert spaces associated with the excited string modes. Composite operators become well defined, *i.e.* have finite matrix elements between arbitrary states in $\mathcal{F}$, through normal ordering. Both the kinematical and the Hamiltonian constraint are hermitean operators. With the normal ordered energy momentum tensors $T_{\pm\pm}(\sigma) \equiv :\big(H(\sigma) \pm D(\sigma)\big):$, the classical constraint algebra is modified by the anomaly (the central term), which is responsible for almost everything that is non-trivial about string theory,

$$\big[T_{\pm\pm}(\sigma)\,,\, T_{\pm\pm}(\sigma')\big] = \partial_\pm\delta(\sigma,\sigma')\big(T_{\pm\pm}(\sigma) + T_{\pm\pm}(\sigma')\big) + \hbar c\,\partial_\pm^3\delta(\sigma,\sigma')\,. \tag{5.5}$$

Our main point in repeating this well known story here is that this algebra holds on the full space of states, *i.e.* prior to the imposition of any constraints, hence closes *off-shell*. Only at this point can we impose and solve the quantum constraints to obtain the subspace $\mathcal{F}_0 \subset \mathcal{F}$ of *physical states* — with the well known result ('No Ghost Theorem') that negative norm states are absent if the number of scalar fields is ≤ 26. Because the Hilbert spaces $\mathcal{F}_0$ and $\mathcal{F}$ of physical and unphysical states, respectively, are well defined (modulo the usual caveats concerning relativistic one-particle wave functions), there is no need to discuss 'habitats' here. The question whether LQG can reproduce these results is currently subject of an intense debate, see [109, 110, 111, 112].

Of course, one key simplifying feature of the string model must be pointed out here, which is not shared by (matter coupled) gravity in higher dimensions: (5.5) is an honest (albeit infinite

[14] Or *'perennials'* in the terminology of [16].

dimensional) Lie algebra. This Lie algebra furthermore possesses a triangular decomposition, allowing for a systematic study of its representations, and this fact is a crucial ingredient in establishing the above results.

5.2 On-shell vs. off-shell closure

We have seen that LQG treats the diffeomorphism and Hamiltonian constraints in a very different manner. Diffeomorphism invariance is implemented by an averaging procedure that makes the states invariant under *finite* (*i.e.* 'exponentiated') diffeomorphisms. In LQG a canonical generator of diffeomorphisms simply does not exist, since the lack of weak continuity that goes with the scalar product (3.7) prevents us from differentiating finite diffeomorphisms so as to obtain 'infinitesimal' ones. By contrast, at least so far, group averaging is of no use, even at the formal level and allowing for whatever approximation scheme, when it comes to imposing the Hamiltonian constraint. Therefore the latter, whose cumbersome form we have exhibited in the foregoing section, must be analysed 'the traditional way'. This restriction applies also to the commutator of two Hamiltonian constraint operators. A further difficulty is that, even for the classical theory, this commutator does not generate an ordinary Lie algebra, in the sense that there appear field dependent structure 'constants', which immediately lead to ordering problems and ambiguities in the constraint algebra. Nevertheless, it should be possible at least to evaluate this commutator as an operatorial expression.

Let us therefore clarify the various notions of closure that can arise: there are at least three different possibilities. As we explained in the foregoing section, the 'true' Hamiltonian $\hat{H}^*$ is defined as a weak limit on some physically motivated subspace of $\mathcal{S}^*$ by means of (4.9) (but, let us repeat, there seems to be no complete consensus in the community as to precisely which subspace to choose). Assuming that this subspace is preserved by the action of $\hat{H}^*$, the strongest notion is 'off-shell closure' (or 'strong closure'), where one seeks to calculate

$$\big[\hat{H}^*[N_1]\,,\,\hat{H}^*[N_2]\big] = \hat{O}^*(N_1;N_2) \tag{5.6}$$

without further restrictions on the states, on which this relation is supposed to hold. The operator $\hat{O}^*(N_1;N_2)$ here is the quantum analogue of the r.h.s. of (5.1), and would be proportional to the diffeomorphism generator. Strong closure is realised for the string model, with the well known result (5.5). If one ignores ordering ambiguities and singular factors, it is also the type of closure expected to work in conventional geometrodynamics. On the other hand, although the calculation in principle does make sense in LQG, too, one is faced with the paradoxical situation that the r.h.s. of (5.6) should not even exist as an operator, as we explained above. While the diffeomorphism generator does not make sense as an operator, it is conceivable that a combination like $g^{mn}H_n$ does exist. However, using this operator one would generate an infinite tower of new operators, and one would therefore no longer be discussing the algebra of constraints.

Equation (5.6) can be relaxed substantially by demanding only 'weak closure'

$$\big[\hat{H}^*[N_1]\,,\,\hat{H}^*[N_2]\big]\mathcal{X} = 0 \tag{5.7}$$

for all states $\mathcal{X}$ in a restricted 'habitat' of states subject to the requirement that the Hamiltonian can be applied twice without leaving the 'habitat'. The latter condition is, for instance, met by the 'vertex smooth' states introduced in [30], which are invariant only under special diffeomorphisms leaving the vertices of a spin network invariant, hence contain $\mathcal{H}_{\rm diff}$ as a subspace (but note that from the point of view of the Hamiltonian action, going to the slightly larger habitat of 'vertex-smooth' states does not really represent an improvement; namely, when acting on a spin network state, the Hamiltonian does not shift nodes around. Hence its action on the habitat or on the larger habitat of 'vertex-smooth' states is essentially the same). Although simpler than (5.6), the

verification of (5.7) in practice still requires knowing the action of the Hamiltonian on all states of the 'habitat' [15].

Finally, an even weaker notion of closure ('ultra-weak closure') is obtained by first calculating the commutator of two regulated Hamiltonians (see section 4.2)

$$\left[\hat{H}[N_1,\epsilon]\,,\,\hat{H}[N_2,\epsilon]\right] = \hat{O}(N_1;N_2;\epsilon) \tag{5.8}$$

on the space $\mathcal{S}$ of spin network states, and by requiring

$$\lim_{\epsilon\to 0}\,\langle\mathcal{X}\,|\,\left[\hat{H}[N_1,\epsilon]\,,\,\hat{H}[N_2,\epsilon]\right]\Psi\rangle = 0 \tag{5.9}$$

where $\mathcal{X}$ is as before, and $\Psi\in\mathcal{S}$. When working on $\mathcal{S}$, this closure for expectation values is the best we can expect, because $\lim_{\epsilon\to 0}\hat{O}(N_1,N_2;\epsilon)\Psi$ does not exist as a limit on $\mathcal{S}$ (for different $\epsilon>0$, these states are all mutually orthogonal by (3.7)). We stress that (5.7) and (5.9) are by no means the same: in (5.7) one uses the dual Hamiltonian defined in (4.9) (where the limit $\epsilon\to 0$ has already been taken) and performs the calculation on a subspace of $\mathcal{S}^*$, whereas the calculation of the commutator in (5.9) takes place on the space $\mathcal{S}$, and the limit $\epsilon\to 0$ is taken only *after* computing the commutator of two regulated Hamiltonians on $\mathcal{S}$. In other words, these two operations (taking the limit $\epsilon\to 0$, and calculating the commutator) need not commute. Because with both (5.7) and (5.9), one forgoes the aim of finding an operatorial expression for the commutator $\left[\hat{H}^*[N_1],\hat{H}^*[N_2]\right]$, making partial use of the constraints, we can say (in a partly supergravity inspired terminology) that the algebra closes 'on-shell'.

Now, direct evaluation of the commutator of two regulated Hamiltonians in (5.9) on a given spin network state, along the lines of section 4.2, would result in a state which is even more complicated than the expressions displayed there. Still, it should be possible in principle to extract from this computation an explicit operatorial expression for the commutator, whose exponentiation would yield a finite diffeomorphism, and thereby ascertain the intrinsic consistency of the theory. Unfortunately, this computation appears to be impossible in practice, so it appears that one is only left with the option of checking weak or ultraweak closure of the algebra on a restricted 'habitat' of states, *i.e.* a physically motivated subspace of $\mathcal{S}^*$ [30]. If it does, one concludes that the algebra closes 'without anomalies'. This conclusion was drawn first in a computation [29, 60] where the habitat was taken to be the Hilbert space $\mathcal{H}_{\rm diff}$ of diffeomorphism invariant states (which is, however, not preserved by the action of the Hamiltonian operator). Later, the constraint algebra was analysed with the same result on a larger habitat of 'vertex-smooth' states [30].

At this point we should like to remark that explicit computations done so far make very little use of the detailed structure of the terms in the Hamiltonian, and also seem to be insensitive to the ambiguities apparent in the latter. Even the technically most advanced investigation of the algebra available so far [30] uses only schematic formulas like (4.22) rather than *explicit* operatorial expressions for the Hamiltonian. On the other hand, computations of more complicated gravity-matter systems (such as the supergravity constraint algebra studied in [58]) show that the detailed structure of all terms in the constraints is indispensable already for the proper closure of the classical constraint algebra. This may indicate that crucial information is lost in restricting the calculation to special states only.

The procedure of making partial use of the constraints *prior* to the evaluation of the constraint algebra is very different from the conventional treatment of quantised canonical systems [44], as

[15] In principle we might even try to include a diffeomorphism averaging into the very definition of the dual Hamiltonian H^*, so as to enforce the preservation of $\mathcal{H}_{\rm diff}$ under the action of $\hat{H}^*$, in which case the absence of anomalies would follow almost by definition.

exemplified by the bosonic string in the previous subsection. There, one first studies the constraint operators and tries to make them well defined as quantum operators without constraining the allowed states, and *before* checking the closure of the constraint algebra. When one then in a second step computes the algebra, one usually encounters subtle quantum (*i.e.* $\hbar$-dependent) modifications of this algebra, whose determination requires great care. Because these modifications are subject to certain constraints (Wess-Zumino type consistency conditions and their offspring in BRST cohomology), they can often be determined in a representation and gauge independent way, as is the case for the group Diff(S^1) (but not for Diff(Σ), unfortunately). If one succeeds in working out this algebra and demonstrating its consistency, the quantum constraint algebra closes 'off-shell'. Only *after* ensuring the closure of the quantum constraint algebra does one proceed to impose and solve the quantum constraints. Let us also remind readers that, at least according to conventional (quantum field theory) wisdom, 'on-shell' calculations of the algebra might be 'empty'. From the very beginning one restricts oneself to diffeomorphism-invariant states, although one does not know whether this symmetry is anomalous or not. In general it is not correct to use a symmetry before it has been shown that it can be implemented without anomalies [113].

Why do we emphasise off-shell closure, despite the difficulties (ordering ambiguities, field dependent structure constants, and the like), which likewise plague other approaches to canonical quantum gravity? Our main reason for emphasising this requirement is that it furnishes an excellent means to distinguish between the 'correct' theory and a mere regularisation: consider for instance replacing the kinetic part of the WDW operator by a smeared point-split operator

$$\frac{\delta}{\delta {e_m}^a(\mathbf{x})}\frac{\delta}{\delta {e_n}^b(\mathbf{x})} \;\longrightarrow\; \int \mathrm{d}\mu(y)\frac{\delta}{\delta {e_m}^a(\mathbf{x}+\mathbf{y})}\frac{\delta}{\delta {e_n}^b(\mathbf{x}-\mathbf{y})} \tag{5.10}$$

with some averaging measure $\mathrm{d}\mu$. With this modification, all constraint operators become well defined, so one might now contemplate 'defining' the quantised theory in terms of them. However, it is immediately clear that the resulting WDW operator is 'wrong': it does not satisfy the correct algebra. As another example, consider modifying the Hamiltonian constraint of string theory by multiplying it with an operator which commutes with all Virasoro generators,

$$(\hat{T}_{++}+\hat{T}_{--}) \;\longrightarrow\; (\hat{T}_{++}+\hat{T}_{--})\,\hat{C}\,, \tag{5.11}$$

where $[\hat{C},\hat{T}_{++}]=0$ and $[\hat{C},\hat{T}_{--}]=0$, while keeping the diffeomorphism constraint $(\hat{T}_{++}-\hat{T}_{--})$ unchanged; for $\hat{C}$ one can for example choose a target-space Lorentz generator. Imposing $(\hat{T}_{++}+\hat{T}_{--})\,\hat{C}|\mathrm{phys}\rangle=0$ would again produce the 'wrong' spectrum, but the algebra would still seem to close on the 'habitat' of physical states annihilated by the diffeomorphism constraint. Thus, if we only demand on-shell closure as in LQG, there is no way of telling whether or not the vanishing of a commutator is merely accidental, that is, not really due to the diffeomorphism invariance of the state, but caused by some other circumstance. This is our main criticism: it appears very likely that the LQG Hamiltonian admits many further modifications on top of the ones we have already discussed, for which the commutator continues to vanish on a suitably restricted habitat of states. Clearly, not all of these can be 'correct'. LQG practitioners who disagree with these statements are hereby challenged to prove them wrong!

6 Conclusions

String theory and LQG pursue the same goal, a consistent theory of quantum gravity, though with very different means. Both approaches address core issues of quantum gravity, but concentrate on complementary aspects of the problem, and have led to valuable insights. For this reason, the

opinion has been voiced that string theory and LQG may ultimately merge together in a grand synthesis, or that LQG might become part of string theory [114, 115, 116][16]. However, on the basis of the available evidence (summarised in this review), this does not appear a likely outcome to us.

We have reported here on the status of loop quantum gravity from a somewhat uncommon perspective. Our general conclusion is that, despite the optimism prevalent in many other reviews, more attention should be paid to basic aspects and unresolved problems of the theory. The large number of ambiguities present in the formalism is, in our opinion, a reason for concern. These ambiguities are most clearly visible in the discussion of the Hamiltonian constraint. As we have stressed, it is this constraint which reflects the main problems of quantum gravity, not only in the loop approach but also in geometrodynamics. However, in the latter approach, it is at least possible to carry out some elementary checks at the dynamical level, and to demonstrate by explicit computation its failure to produce a theory of quantum gravity with all the requisite properties. For the time being, such an *examen crucis* does not appear possible for LQG due to the mathematical complexity of the Hamiltonian constraint, despite many other proposed tests of its viability in the literature. Another important issue which has not received the attention which it deserves is the ultra-local behaviour of the Hamiltonian constraint in LQG. Without an understanding of this problem, it seems premature to discuss applications of the formalism, which assume that a classical limit exists, or possible experimental signatures.

In summary, we feel that there are still too many problems at a basic level that need to be addressed and resolved before one can tell whether the loop quantum gravity programme will succeed or fail. We hope that the present paper will provide an incentive to re-focus attention to these basic issues.

Acknowledgements

This work grew out of the Workshop "Strings meet Loops" at the AEI in October 2003, and the numerous discussions that followed it. We would especially like to thank Abhay Ashtekar, Martin Bojowald, Robert Helling, Marc Henneaux, Jurek Lewandowski, Renate Loll, Dieter Maison, Don Marolf, Heide Narnhofer, Hendryk Pfeiffer, Jan Plefka, Nicolai Reshetikhin, Ingo Runkel, Urs Schreiber, Aureliano Skirzewski, Matthias Staudacher and Walter Thirring for useful discussions and correspondence. Special thanks go to Kirill Krasnov and Thomas Thiemann for insightful comments on the manuscript, and especially to Thomas Thiemann for patiently putting up with our criticism.

[16] Finding a connection between *classical* two-form gravity in four dimensions and *topological* string theory [115] can hardly be considered evidence for the connection between LQG and string theory. By this logic, any theory which starts from classical gravity would be connected to string theory.

References

[1] M. H. Goroff and A. Sagnotti, "Quantum gravity at two loops", *Phys. Lett.* **B160** (1985) 81.

[2] M. H. Goroff and A. Sagnotti, "The ultraviolet behavior of Einstein gravity", *Nucl. Phys.* **B266** (1986) 709.

[3] A. E. M. van de Ven, "Two loop quantum gravity", *Nucl. Phys.* **B378** (1992) 309–366.

[4] S. Weinberg, "Ultraviolet divergences in quantum gravity", in "An Einstein centenary survey", S. Hawking and W. Israel, eds., pp. 790–832. Cambridge University Press, 1979.

[5] S. Weinberg, "What is quantum field theory, and what did we think it was?", `hep-th/9702027`.

[6] O. Lauscher and M. Reuter, "Is quantum Einstein gravity nonperturbatively renormalizable?", *Class. Quant. Grav.* **19** (2002) 483–492, `hep-th/0110021`.

[7] G. W. Gibbons and S. W. Hawking, "Action integrals and partition functions in quantum gravity", *Phys. Rev.* **D15** (1977) 2752–2756.

[8] S. W. Hawking, "The path-integral approach to quantum gravity", in "An Einstein centenary survey", S. Hawking and W. Israel, eds., pp. 746–789. Cambridge University Press, 1979.

[9] J. B. Hartle and S. W. Hawking, "Wave function of the universe", *Phys. Rev.* **D28** (1983) 2960–2975.

[10] R. Gambini and J. Pullin, "Loops, knots, gauge theories and quantum gravity", Cambridge University Press, 1996.

[11] T. Thiemann, "Introduction to modern canonical quantum general relativity", `gr-qc/0110034`.

[12] A. Ashtekar and J. Lewandowski, "Background independent quantum gravity: A status report", *Class. Quant. Grav.* **21** (2004) R53, `gr-qc/0404018`.

[13] C. Rovelli, "Quantum gravity", Cambridge Monographs on Mathematical Physics, 2004.

[14] A. Perez, "Introduction to loop quantum gravity and spin foams", `gr-qc/0409061`.

[15] G. T. Horowitz, "Ashtekar's approach to quantum gravity", in "Strings and symmetries", pp. 154–167. Stony Brook, 1991. `hep-th/9109002`.

[16] K. V. Kuchar, "Canonical quantum gravity", `gr-qc/9304012`.

[17] S. Carlip, "Quantum gravity: A progress report", *Rep. Prog. Phys.* **64** (2001) 885, `gr-qc/0108040`.

[18] H. Zeh, "The physical basis of the direction of time, 3rd edition", Springer Verlag, 1999.

[19] C. Isham, "Conceptual and geometrical problems in quantum gravity", in "Recent Aspects of Quantum Fields", H. Mitter and H. Gausterer, eds. Springer Lecture Notes in Physics 396, 1991.

[20] T. Banks, "T C P, quantum gravity, the cosmological constant and all that...", *Nucl. Phys.* **B249** (1985) 332.

[21] A. Ashtekar, S. Fairhurst, and J. L. Willis, "Quantum gravity, shadow states, and quantum mechanics", *Class. Quant. Grav.* **20** (2003) 1031–1062, `gr-qc/0207106`.

[22] T. Thiemann, "Gauge field theory coherent states (GCS). I: General properties", *Class. Quant. Grav.* **18** (2001) 2025–2064, `hep-th/0005233`.

[23] T. Thiemann and O. Winkler, "Gauge field theory coherent states (GCS). II: Peakedness properties", *Class. Quant. Grav.* **18** (2001) 2561–2636, `hep-th/0005237`.

[24] T. Thiemann and O. Winkler, "Gauge field theory coherent states (GCS) III: Ehrenfest theorems", *Class. Quant. Grav.* **18** (2001) 4629–4682, `hep-th/0005234`.

[25] T. Thiemann and O. Winkler, "Gauge field theory coherent states (GCS). IV: Infinite tensor product and thermodynamical limit", *Class. Quant. Grav.* **18** (2001) 4997–5054, `hep-th/0005235`.

[26] T. Thiemann, "Complexifier coherent states for quantum general relativity", `gr-qc/0206037`.

[27] H. Sahlmann, T. Thiemann, and O. Winkler, "Coherent states for canonical quantum general relativity and the infinite tensor product extension", *Nucl. Phys.* **B606** (2001) 401–440, `gr-qc/0102038`.

[28] T. Thiemann, "The Phoenix project: Master constraint programme for loop quantum gravity", `gr-qc/0305080`.

[29] T. Thiemann, "Anomaly-free formulation of non-perturbative, four-dimensional Lorentzian quantum gravity", *Phys. Lett.* **B380** (1996) 257–264, gr-qc/9606088.

[30] J. Lewandowski and D. Marolf, "Loop constraints: A habitat and their algebra", *Int. J. Mod. Phys.* **D7** (1998) 299–330, gr-qc/9710016.

[31] M. Bojowald, "Absence of singularity in loop quantum cosmology", *Phys. Rev. Lett.* **86** (2001) 5227–5230, gr-qc/0102069.

[32] A. Ashtekar, M. Bojowald, and J. Lewandowski, "Mathematical structure of loop quantum cosmology", *Adv. Theor. Math. Phys.* **7** (2003) 233–268, gr-qc/0304074.

[33] M. Bojowald, "Inflation from quantum geometry", *Phys. Rev. Lett.* **89** (2002) 261301, gr-qc/0206054.

[34] J. D. Bekenstein, "Black holes and entropy", *Phys. Rev.* **D7** (1973) 2333–2346.

[35] S. W. Hawking, "Particle creation by black holes", *Commun. Math. Phys.* **43** (1975) 199–220.

[36] A. Ashtekar, J. Baez, A. Corichi, and K. Krasnov, "Quantum geometry and black hole entropy", *Phys. Rev. Lett.* **80** (1998) 904–907, gr-qc/9710007.

[37] A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy", *Phys. Lett.* **B379** (1996) 99–104, hep-th/9601029.

[38] M. P. Reisenberger, "World sheet formulations of gauge theories and gravity", gr-qc/9412035.

[39] M. P. Reisenberger and C. Rovelli, "'Sum over surfaces' form of loop quantum gravity", *Phys. Rev.* **D56** (1997) 3490–3508, gr-qc/9612035.

[40] J. C. Baez, "Spin foam models", *Class. Quant. Grav.* **15** (1998) 1827–1858, gr-qc/9709052.

[41] R. Arnowitt, S. Deser, and C. W. Misner, "The dynamics of general relativity", in "Gravitation: an introduction to current research", L. Witten, ed., pp. 227–265. Wiley, 1962. gr-qc/0405109.

[42] C. Misner, K. Thorne, and J. Wheeler, "Gravitation", Freeman, 1973.

[43] P. A. M. Dirac, "General theory of relativity", Princeton University Press, 1975.

[44] M. Henneaux and C. Teitelboim, "Quantization of gauge systems", Princeton University Press, 1992.

[45] C. Kiefer, "Quantum Gravity", Clarendon Press, 2004.

[46] H. Nicolai and H. J. Matschull, "Aspects of canonical gravity and supergravity", *J. Geom. Phys.* **11** (1993) 15–62.

[47] B. S. DeWitt, "Quantum theory of gravity. I. The canonical theory", *Phys. Rev.* **160** (1967) 1113–1148.

[48] B. S. DeWitt, "Quantum theory of gravity. II. The manifestly covariant theory", *Phys. Rev.* **162** (1967) 1195–1239.

[49] B. S. DeWitt, "Quantum theory of gravity. III. Applications of the covariant theory", *Phys. Rev.* **162** (1967) 1239–1256.

[50] J. A. Wheeler, "Geometrodynamics and the issue of the final state", in "Relativity, Groups and Topology", C. DeWitt and J. Wheeler, eds. Blackie and Sons, London, 1964.

[51] A. Ashtekar, "New Hamiltonian formulation of general relativity", *Phys. Rev.* **D36** (1987) 1587–1602.

[52] S. Holst, "Barbero's Hamiltonian derived from a generalized Hilbert-Palatini action", *Phys. Rev.* **D53** (1996) 5966–5969, gr-qc/9511026.

[53] J. Samuel, "Comment on Holst's Lagrangian formulation", *Phys. Rev.* **D63** (2001) 068501.

[54] J. F. Barbero, "Real Ashtekar variables for Lorentzian signature space times", *Phys. Rev.* **D51** (1995) 5507–5510, gr-qc/9410014.

[55] G. Immirzi, "Real and complex connections for canonical gravity", *Class. Quant. Grav.* **14** (1997) L177–L181, gr-qc/9612030.

[56] R. Gambini, O. Obregon, and J. Pullin, "Yang-Mills analogues of the Immirzi ambiguity", *Phys. Rev.* **D59** (1999) 047505, gr-qc/9801055.

[57] I. Bengtsson, "Yang-Mills theory and general relativity in three and four dimensions", *Phys. Lett.* **B220** (1989) 51.

[58] H. Nicolai, "The canonical structure of maximally extended supergravity in three dimensions", *Nucl. Phys.* **B353** (1991) 493–518.

[59] H. J. Matschull and H. Nicolai, "Canonical quantum supergravity in three dimensions", *Nucl. Phys.* **B411** (1994) 609–646, gr-qc/9306018.

[60] T. Thiemann, "Quantum spin dynamics (QSD)", *Class. Quant. Grav.* **15** (1998) 839–873, gr-qc/9606089.

[61] J. Samuel, "Is Barbero's Hamiltonian formulation a gauge theory of Lorentzian gravity?", *Class. Quant. Grav.* **17** (2000) L141–L148, gr-qc/0005095.

[62] H. Kodama, "Holomorphic wave function of the universe", *Phys. Rev.* **D42** (1990) 2548–2565.

[63] R. Jackiw, "Topological investigations in quantized gauge theories", in "Current algebra and anomalies", S. B. Treiman, ed., pp. 258, exercise 3.7. World Scientific, 1985.

[64] E. Witten, "A note on the Chern-Simons and Kodama wavefunctions", gr-qc/0306083.

[65] L. Freidel and L. Smolin, "The linearization of the Kodama state", *Class. Quant. Grav.* **21** (2004) 3831–3844, hep-th/0310224.

[66] C. Rovelli and L. Smolin, "Loop space representation of quantum general relativity", *Nucl. Phys.* **B331** (1990) 80.

[67] B. Simon, "The $P(\phi)_2$ Euclidean (quantum) field theory", Princeton University Press, 1974. Princeton Series in Physics.

[68] A. Ashtekar, A. Corichi, and J. A. Zapata, "Quantum theory of geometry. III: Non-commutativity of Riemannian structures", *Class. Quant. Grav.* **15** (1998) 2955–2972, gr-qc/9806041.

[69] M. H. Stone, "Linear transformations in Hilbert space, III: operational methods and group theory", *Proc. Nat. Acad. Sci.* **16** (1930) 172–175.

[70] J. von Neumann, "Die Eindeutigkeit der Schrödingerschen Operatoren", *Math. Ann.* **104** (1931) 570.

[71] J. Rosenberg, "A selective history of the Stone-von Neumann theorem", in "Operator algebras, quantization, and noncommutative geometry: a centennial celebration honoring John von Neumann and Marshall H. Stone", R. S. Doran and R. V. Kadison, eds., vol. 365 of *Contemporary Mathematics*.

[72] A. Perez, "Spin foam models for quantum gravity", *Class. Quant. Grav.* **20** (2003) R43, gr-qc/0301113.

[73] W. Thirring and H. Narnhofer, "Covariant QED without indefinite metric", *Rev. Math. Phys.* **SI1** (1992) 197–211.

[74] A. A. Migdal, "Loop equations and $1/N$ expansion", *Phys. Rev.* **102** (1983) 199–290.

[75] A. M. Polyakov, "Gauge fields and strings", Harwood, 1987. Contemporary concepts in physics, volume 3.

[76] I. Montvay and G. Münster, "Quantum fields on a lattice", Cambridge University Press, 1994.

[77] H. J. Rothe, "Lattice gauge theories; an introduction", World Scientific, 1992.

[78] M. Creutz, "Quarks, gluons and lattices", Cambridge University Press, 1983.

[79] B. Dittrich, "Partial and complete observables for hamiltonian constrained systems", gr-qc/0411013.

[80] T. Thiemann, "Reduced phase space quantization and dirac observables", gr-qc/0411031.

[81] A. Ashtekar and J. Lewandowski, "Quantum theory of geometry. I: Area operators", *Class. Quant. Grav.* **14** (1997) A55–A82, gr-qc/9602046.

[82] A. Ashtekar and J. Lewandowski, "Quantum theory of geometry. II: Volume operators", *Adv. Theor. Math. Phys.* **1** (1998) 388, gr-qc/9711031.

[83] M. Domagala and J. Lewandowski, "Black hole entropy from quantum geometry", *Class. Quant. Grav.* **21** (2004) 5233–5244, gr-qc/0407051.

[84] K. A. Meissner, "Black hole entropy in loop quantum gravity", *Class. Quant. Grav.* **21** (2004) 5245–5252, gr-qc/0407052.

[85] A. Ghosh and P. Mitra, "An improved lower bound on black hole entropy in the quantum geometry approach", gr-qc/0411035.

[86] T. Thiemann, "Closed formula for the matrix elements of the volume operator in canonical quantum gravity", *J. Math. Phys.* **39** (1998) 3347–3371, gr-qc/9606091.

[87] J. Brunnemann and T. Thiemann, "Simplification of the spectral analysis of the volume operator in loop quantum gravity", gr-qc/0405060.

[88] J. Glimm and A. Jaffe, "Quantum physics – A functional integral point of view", Springer, New York, 1987.

[89] J. C. Baez and K. V. Krasnov, "Quantization of diffeomorphism-invariant theories with fermions", *J. Math. Phys.* **39** (1998) 1251–1271, hep-th/9703112.

[90] T. Thiemann, "QSD V: Quantum gravity as the natural regulator of matter quantum field theories", *Class. Quant. Grav.* **15** (1998) 1281–1314, gr-qc/9705019.

[91] M. Varadarajan, "Fock representations from U(1) holonomy algebras", *Phys. Rev.* **D61** (2000) 104001, gr-qc/0001050.

[92] M. Varadarajan, "Photons from quantized electric flux representations", *Phys. Rev.* **D64** (2001) 104003, gr-qc/0104051.

[93] A. Ashtekar and J. Lewandowski, "Relation between polymer and Fock excitations", *Class. Quant. Grav.* **18** (2001) L117–L128, gr-qc/0107043.

[94] A. Ashtekar, J. Lewandowski, and H. Sahlmann, "Polymer and Fock representations for a scalar field", *Class. Quant. Grav.* **20** (2003) L11–1, gr-qc/0211012.

[95] T. Thiemann, "Kinematical Hilbert spaces for fermionic and Higgs quantum field theories", *Class. Quant. Grav.* **15** (1998) 1487–1512, gr-qc/9705021.

[96] P. Peldan, "Ashtekar's variables for arbitrary gauge group", *Phys. Rev.* **D46** (1992) 2279–2282, hep-th/9204069.

[97] S. Chakraborty and P. Peldan, "Towards a unification of gravity and Yang-Mills theory", *Phys. Rev. Lett.* **73** (1994) 1195–1198, gr-qc/9401028.

[98] S. Melosch and H. Nicolai, "New canonical variables for $D = 11$ supergravity", *Phys. Lett.* **B416** (1998) 91–100, hep-th/9709227.

[99] P. Hajicek and K. V. Kuchar, "Constraint quantization of parametrized relativistic gauge systems in curved space-times", *Phys. Rev.* **D41** (1990) 1091–1104.

[100] B. de Wit, H. J. Matschull, and H. Nicolai, "Physical states in $d = 3$, $N = 2$ supergravity", *Phys. Lett.* **B318** (1993) 115–121, gr-qc/9309006.

[101] J. A. Zapata, "A combinatorial approach to diffeomorphism invariant quantum gauge theories", *J. Math. Phys.* **38** (1997) 5663–5681, gr-qc/9703037.

[102] W. Fairbairn and C. Rovelli, "Separable Hilbert space in loop quantum gravity", *J. Math. Phys.* **45** (2004) 2802–2814, gr-qc/0403047.

[103] L. Smolin, "The classical limit and the form of the Hamiltonian constraint in non-perturbative quantum general relativity", gr-qc/9609034.

[104] D. E. Neville, "Long range correlations in quantum gravity", *Phys. Rev.* **D59** (1999) 044032, gr-qc/9803066.

[105] R. Loll, "On the diffeomorphism-commutators of lattice quantum gravity", *Class. Quant. Grav.* **15** (1998) 799–809, gr-qc/9708025.

[106] I. Kouletsis and K. V. Kuchar, "Diffeomorphisms as symplectomorphisms in history phase space: Bosonic string model", *Phys. Rev.* **D65** (2002) 125026, gr-qc/0108022.

[107] M. B. Green, J. H. Schwarz, and E. Witten, "Superstring theory", Cambridge University Press, 1987.

[108] J. Polchinski, "String theory", Cambridge University Press, 1998.

[109] A. Starodubtsev, "String theory in a vertex operator representation: A simple model for testing loop quantum gravity", gr-qc/0201089.

[110] T. Thiemann, "The LQG string: Loop quantum gravity quantization of string theory. I: Flat target space", hep-th/0401172.

[111] U. Schreiber, "DDF and Pohlmeyer invariants of (super)string", *JHEP* **05** (2004) 027, hep-th/0403260.

[112] R. C. Helling and G. Policastro, "String quantization: Fock vs. LQG representations", hep-th/0409182.

[113] F. De Jonghe, R. Siebelink, W. Troost, S. Vandoren, P. van Nieuwenhuizen, and A. Van Proeyen, "The regularized BRST jacobian of pure Yang-Mills theory", *Phys. Lett.* **B289** (1992) 354–360, hep-th/9206097.

[114] G. Musser, "The future of string theory - A conversation with Brian Greene", *Sci. Am.*, November 2003 48–53.

[115] R. Dijkgraaf, S. Gukov, A. Neitzke, and C. Vafa, "Topological M-theory as unification of form theories of gravity", hep-th/0411073.

[116] D. Overbye, "String theory, at 20, explains it all (or not)", *New York Times*, December 7 2004.