Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Betreuer

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Kinematic Structure and Diffeomorphism Invariance in Loop Quantum Cosmology

ausgeführt am Institut für
Theoretische Physik
der Technischen Universität Wien

unter Anleitung von
Privatdoz. Dipl. Ing. Dr techn. Herbert Balasin

durch

Thomas Mörtlbauer
Volksgartenstraße 6
A-4020 Linz

Ort, Datum

Unterschrift

Acknowledgments

First of all I would like to thank my supervisor Prof. Herbert Balasin for his guidance and his open ear for all of my questions and problems. Without his support and continued belief in this project I would probably not have found the strength and persistence to finish this work. I am also indebted to him for his inspiring lessons on general relativity and for raising my interest for the promising subject of loop quantum gravity in the first place.

I could not have completed this thesis without the support from my parents. Not only did they provide financial support. From the beginning they encouraged my growing interest in understanding how the universe "works". Without reservation they allowed me to pursue my interests and find my path from physics to mathematics and back to theoretical physics in an attempt to find answers to the most fundamental questions of physics. I thank them and my sister Lisa for making me the person I am today.

I am grateful to the Institute for Theoretical Physics at the Vienna University of Technology for supplying a workplace. This not only provided a quiet place to write my thesis but also gave me the opportunity to meet new friends and colleagues. Especially I want to mention my former roommate Wolfgang Wieland, who patiently answered my questions when I took my first steps in loop quantum gravity and Florian Kren — we mastered most of the mathematics studies together and especially enjoyed our mutual interest in functional analysis.

A special thanks goes to the most important person in my life — my beloved wife Kathrin. Her exceptional love and support and her unconditional belief in me kept me going through times when I had doubts with regard to my professional future.

Contents

Acknowledgments i

Notation, conventions and some formulas v

Introduction ix

Part 1. Loop Quantum Gravity 1

Chapter 1. Hamiltonian Formulation of general relativity 3
1. Nature of the initial value formulation of general relativity 4
2. Canonical form of the action 7
3. Legendre transform 9
4. Motions generated by the constraints 12
5. Hypersurface deformation algebra 13

Chapter 2. Gauge theory formulation of general relativity 17
1. The quantization program 17
2. Foundations of gauge theories 20
3. New variables for general relativity 23
4. Equivalence of gauge theory- and ADM-formulation 26

Chapter 3. Loop quantum gravity 31
1. The holonomy-flux algebra 31
2. Representations 39
3. Implementation of the kinematical constraints 48
4. Kinematical geometrical operators 54
5. Implementation of the Hamiltonian constraint 58

Part 2. Loop Quantum Cosmology 67

Overview 69
Cosmological Models 70

Chapter 4. Symmetry Reduction 73
1. Symmetry Reduction for Theories of Connections 73
2. Bianchi Models in the Connection Framework 78

Chapter 5. Original Approach to Loop Quantum Cosmology 81

1. Kinematical Hilbert Spaces 81
2. Gauß Constraint 84
3. Diffeomorphism Constraint 86
4. Diffeomorphism invariant states for isotropic models 89
5. Volume Operator in isotropic LQC 92
6. Quantization of the Hamiltonian Constraint 96

Chapter 6. Simplified Kinematics 103
1. Improvements for the symmetry-reduced phase space 103
2. The cosmological equivalent of the holonomy-flux algebra 104
3. Loop representation of the holonomy-flux algebra 106
4. Quantization of the Hamiltonian constraint 112

Chapter 7. Discussion and Comparison 115
1. Interesting results of loop quantum cosmology 115
2. Comparison of both approaches 117
3. Suggestions for improvements of the kinematical structure 118

Conclusion and Outlook 125

Part 3. Appendices 127

Appendix A. Hypersurfaces 129
1. Definitions 129
2. Extrinsic geometry 130
3. Intrinsic geometry 132
4. Stokes's theorem 136

Appendix B. Lagrangian Formulation of General Relativity 139
1. Preliminaries 139
2. Hilbert-term 141
3. Hawking boundary term 142

Appendix C. Symplectic geometry 145

Appendix D. Mathematical supplements for physicists 149
1. Representation theory 149
2. Miscellaneous 153

Bibliography 155

Notation, conventions and some formulas

- We use the sign conventions of Misner, Thorne and Wheeler [28]. In particular, our metric signature is $-+++$, which has the advantage of inducing a positive definite metric on spacelike hypersurfaces.
- We use latin letters from the beginning of the alphabet $a, b, \ldots$ (and $m, n, \ldots$ when we run out of those) to denote abstract spacetime and spatial indices, and greek letters $\mu, \nu, \ldots$ for their components. When it might not be clear from the context, we will use a prefix to denote three-dimensional spatial tensors, as in ${}^3R^a{}_{bcd}$. In that case, tensors without prefix will be four-dimensional objects.

  Latin letters from the middle of the alphabet will denote internal indices. We use capital letters $I, J, \ldots$ for tetrad components and lower case letters $i, j, \ldots$ for triad components.

  To indicate tensor densities, we will use a tilde, as in $\widetilde{V}_a$. Sometimes we use multiple tildes to represent tensor densities of weight greater than one, and we will write the tilde under the tensor symbol to denote negative weights.
- The gravitational coupling constant is $\kappa = \frac{16\pi G}{c^3}$ and the Planck length is $l_P = \sqrt{\hbar\kappa}$. Most of the time we will set κ and $\hbar$ (and therefore also l_P) to unity and only reintroduce them when we are interested in numerical values. In most literature on LQG it is customary to redefine the gravitational constant as $8\pi G$ in order to get rid of the factor $\frac{1}{2}$ in the Poisson brackets (2.23). Alternatively some authors also redefine $\widetilde{E}^a_i$ to achieve the same effect. We choose not do either at the price of introducing a factor $\frac{1}{2}$ as compared to the literature in some formulas.
- All derivative operators are defined to only act on the quantity right next to them. In particular this applies not only to covariant derivatives ∇_a, but also to Lie derivatives and variations.
- The four-dimensional spacetime metric is denoted by g_{ab}, the spatial metric by h_{ab}. When we need the flat Minkowski-metric we denote it by η_{ab}.
- We will often omit indices that are contracted, as for example in $(n\nabla) := n^a\nabla_a$, $K := h^{ab}K_{ab}$. Note that for tangential objects we can equally well contract with h_{ab} or g_{ab}.

- Symmetrizing and antisymmetrizing is denoted respectively by

$$t_{(a_1\dots a_n)} := \frac{1}{n!}\sum_\pi t_{a_{\pi(1)}\dots a_{\pi(n)}}$$
$$t_{[a_1\dots a_n]} := \frac{1}{n!}\sum_\pi \mathrm{sgn}(\pi) t_{a_{\pi(1)}\dots a_{\pi(n)}},$$

  where the sum runs over all permutations of the indices. Indices that are to be excluded from (anti)symmetrizing will be enclosed in vertical bars, as in $t_{[a_1\dots|a_i|\dots a_n]}$.
- Collecting p indices $a_1,\dots a_p$ in a clumped index $I_p = (a_1,\dots a_p)$ provides a convenient way to write general formulas for tensors. We define the symbol

$$\delta^{I_p}_{J_p} := \det\begin{pmatrix}\delta^{a_1}_{b_1} & \dots & \delta^{a_p}_{b_1}\\ \vdots & \ddots & \vdots\\ \delta^{a_1}_{b_p} & \dots & \delta^{a_p}_{b_p}\end{pmatrix} = \begin{cases} 1 & \exists\sigma\in S_p, \sigma(I_p)=J_p, \mathrm{sgn}(\sigma)=1\\ -1 & \exists\sigma\in S_p, \sigma(I_p)=J_p, \mathrm{sgn}(\sigma)=-1\\ 0 & \text{else}\end{cases}. \tag{N.1}$$

  Note that $\frac{1}{p!}\delta^{I_p}_{J_p}$ is the projector on antisymmetric tensors. We can now write the wedge product of forms as

$$(\omega_p\wedge\varphi_q)_{I_{p+q}} := \frac{1}{p!q!}\delta^{I_pI_q}_{I_{p+q}}\omega_{I_p}\varphi_{I_q}.$$

  Exterior differentiation is written with an arbitrary covariant derivative ∇_a as

$$(d\omega)_{I_{p+1}} := \frac{1}{p!}\delta^{aI_p}_{I_{p+1}}\nabla_a\omega_{I_p}.$$

  We denote the inner product of vector fields and forms by $(v\lrcorner\,\omega)_{I_{p-1}} := v^a\omega_{aI_{p-1}}$. Sometimes we will also resort to the notation $i_v\omega$. This notation greatly simplifies the writing of proofs such as the anti-derivation property of exterior differential and interior product,

$$d(\omega\wedge\varphi) = d\omega\wedge\varphi + (-1)^p\omega\wedge d\varphi$$
$$v\lrcorner\,(\omega\wedge\varphi) = (v\lrcorner\,\omega)\wedge\varphi + (-1)^p\omega\wedge(v\lrcorner\,\varphi)$$

  for p-forms ω and q-forms φ.
- The difference-tensor is defined for one-forms and vectors by

$$\begin{aligned}\nabla_a\omega_b &= \widetilde{\nabla}_a\omega_b - C^c{}_{ba}\omega_c\\ \nabla_a v^b &= \widetilde{\nabla}_a v^b + C^b{}_{ca}v^c.\end{aligned} \tag{N.2}$$

  If ∇ is compatible with g_{ab} and $\widetilde{\nabla}$ is compatible with $\widetilde{g}_{ab}$, then it is given explicitly by

$$C^c{}_{ab} = \frac{1}{2}g^{cm}\left(\widetilde{\nabla}_b g_{ma} + \widetilde{\nabla}_a g_{bm} - \widetilde{\nabla}_m g_{ab}\right). \tag{N.3}$$

- At one point we will have to deal with derivations which are not symmetric, i.e. the second derivatives do not commute on scalars. We define the torsion tensor $T^c{}_{ab}$ via

$$[\nabla_a,\nabla_b]\,f =: T^c{}_{ab}\,\nabla_c f. \tag{N.4}$$

Expressing ∇ in terms of an arbitrary symmetric derivative $\widetilde{\nabla}$ and a difference tensor $C^a{}_{bc}$ the torsion tensor can be calculated by

$$T^c{}_{ab} = C^c{}_{ab} - C^c{}_{ba}, \tag{N.5}$$

as can be seen easily with (N.2) and remembering that the first derivatives agree on scalars since the are just given by the directional derivative.

- The Riemann tensor

$$\begin{aligned} [\nabla_a, \nabla_b]\omega_c &= -R^d{}_{cab}\omega_d \\ [\nabla_a, \nabla_b]v^c &= \quad R^c{}_{dab}v^d \end{aligned} \tag{N.6}$$

can be computed in terms of the difference-tensor via

$$R^a{}_{bcd} = \widetilde{\nabla}_c C^a{}_{bd} - \widetilde{\nabla}_d C^a{}_{bc} + C^a{}_{mc}C^m{}_{bd} - C^a{}_{md}C^m{}_{bc} + \widetilde{R}^a{}_{bcd}, \tag{N.7}$$

where $\widetilde{R}^a{}_{bcd}$ is the Riemann tensor corresponding to $\widetilde{\nabla}$. The Ricci tensor is obtained via contraction,

$$R_{ab} := R^c{}_{acb}. \tag{N.8}$$

- The Hodge dual of a p-form ω on an n-dimensional manifold is the $(n-p)$-form

$$(*\omega)_{a_1\ldots a_{n-p}} := \frac{1}{p!}\varepsilon^{b_1\ldots b_p}{}_{a_1\ldots a_{n-p}}\omega_{b_1\ldots b_p}. \tag{N.9}$$

Applying the Hodge star operator again, we get

$$**\omega = (-1)^{s+p(n-p)}\omega, \tag{N.10}$$

s being the signature of the metric.

- Formulas for the Lie derivative:

$$\begin{aligned} \mathcal{L}_v t^{a_1\ldots a_k}{}_{b_1\ldots b_l} &= v^c\nabla_c t^{a_1\ldots a_k}{}_{b_1\ldots b_l} - \sum_{i=1}^{k} t^{a_1\ldots c\ldots a_k}{}_{b_1\ldots b_l}\nabla_c v^{a_i} \\ &\quad + \sum_{j=1}^{l} t^{a_1\ldots a_k}{}_{b_1\ldots c\ldots b_l}\nabla_{b_j} v^c \end{aligned} \tag{N.11}$$

$$\mathcal{L}_v\omega = d\,(v \lrcorner\, \omega) + v \lrcorner\, d\omega \tag{N.12}$$

$$\mathcal{L}_{[v,w]} = [\mathcal{L}_v, \mathcal{L}_w] \tag{N.13}$$

$$[\mathcal{L}_v, i_w] = i_{[v,w]} \tag{N.14}$$

- Since n-forms can also be viewed as scalar densities, it will be convenient to have a notation to switch between these different points of view. We can achieve this by defining the symbols $\widetilde{\delta}^{abc}$ and $\underset{\sim}{\delta}_{abc}$. The first is a tensor density of weight one and is defined via $(\widetilde{\delta}^{abc})_{pqr} := \delta^{abc}_{pqr}$, the definition of the latter is analogous. Their use is illustrated by the identities $\omega_h\underset{\sim}{\delta}_{abc} = \epsilon_{abc}$, $\omega_h = \frac{1}{3!}\epsilon_{abc}\widetilde{\delta}^{abc}$, with the volume form ω_h and the epsilon tensor $\epsilon_{abc} = \sqrt{\det h}(dx^1 \wedge dx^2 \wedge dx^3)_{abc}$ in any oriented coordinate system.

In other words, $\underset{\sim}{\delta}_{abc}$ just adds the 3-form indices of a tensor density, while contraction with $\frac{1}{3!}\widetilde{\delta}^{abc}$ removes indices of a form and creates a density.

- For tensors with values in a given Lie algebra $\mathfrak{g}$ we define a bracket by

$$[t^{a\ldots}{}_{b\ldots}, u^{c\ldots}{}_{d\ldots}] := t^{a\ldots}{}_{b\ldots}{}^{j} u^{c\ldots}{}_{d\ldots}{}^{k} [\tau_j, \tau_k] , \tag{N.15}$$

where τ_i are the generators of $\mathfrak{g}$ and the tensors have been decomposed in this basis, $t^{a\ldots}{}_{b\ldots} = t^{a\ldots}{}_{b\ldots}{}^{j}\tau_j$. If we are dealing with forms instead of tensors, the tensor product is replaced by the wedge product. Furthermore, for forms we have the equivalent definition $[\phi, \psi]^i{}_k = \phi^i{}_j \wedge \psi^j{}_k - (-1)^{pq}\psi^i{}_j\phi^j{}_k$. Since we therefore have two definitions of the bracket, we agree that we take the wedge product of the component forms as long as we are dealing with forms only and the tensor product otherwise. Exceptions will be stated explicitly.

The bracket has the following properties, here outlined for forms ϕ, ψ, θ of degree p, q, r respectively (see e.g. [20]):

$$[\phi_1 + \phi_2, \psi] = [\phi_1, \psi] + [\phi_2, \psi] \tag{N.16}$$

$$[\phi, \psi] = (-1)^{pq+1} [\psi, \phi] \tag{N.17}$$

$$[\phi, [\psi, \theta]] = [[\phi, \psi] , \theta] + (-1)^{pq} [\psi, [\phi, \theta]] \tag{N.18}$$

$$d [\phi, \psi] = [d\phi, \psi] + (-1)^p [\phi, d\psi] . \tag{N.19}$$

For tensors, the sign factors are -1 in the second equation, $+1$ in the third and fourth.

- The determinant of an $n \times n$-matrix M can be written as

$$\det M = \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon^{j_1\ldots j_n} M^{i_1}{}_{j_1} \ldots M^{i_n}{}_{j_n} , \tag{N.20}$$

which follows from the well known Leibniz formula $\det M = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n M^i_{\sigma(i)}$ by writing the sum over all permutations with an ε-symbol and contracting with another ε-symbol.

Introduction

Quantum mechanics and general relativity are the two dominant physical theories of our time. Both are experimentally well-tested, are therefore widely accepted and have led to numerous applications. But their discovery at the beginning of the twentieth century also came with a great conceptual revolution which destroyed the coherent picture of physics that had formed until that time. Quantum mechanics teaches us that any physical quantity is quantized and its time-evolution is governed by probabilistic laws. On the other hand, general relativity is formulated in terms of a smooth metric field, whose evolution is described by the Einstein equations and is therefore deterministic. Vice versa, quantum mechanics is formulated using a fixed time variable. More generally, quantum field theories are formulated with respect to flat Minkowski space. This contradicts the central idea of general relativity, which is that gravity is geometry. Put more explicitly, matter does not simply evolve on a static stage provided by a background Riemannian manifold. In contrast, the metric of this manifold is itself a dynamical entity, the metric and matter fields become equal actors and the stage has disappeared (or has at least been downgraded to a mere differentiable manifold). Thus, both theories fail to implement the key insights gained from the other.

Also both theories suffer from internal inconsistencies. General relativity predicts space-time singularities — in the simplest case these are regions of space-time where curvature or matter-density become infinite. The most prominent examples of these situations are black holes and the initial singularity of cosmology, the "big bang". Penrose and Hawking even proved that these singularities are a generic feature of every space-time, provided some reasonable physical assumptions are met. See [25] for an overview of their famous singularity theorems. Quantum field theory on the other hand suffers from UV-divergencies at short distances.

This short discussion should make the need for a theory apparent, which unites quantum theory and general relativity — a quantum theory of gravity or "quantum gravity". But despite over half a century of continuous research on the subject we still do not have a consistent theory of quantum gravity.

The most straightforward approach would be to split the space-time metric into a (Minkowski) background $\eta_{\mu\nu}$ and a "small" perturbation $h_{\mu\nu}$, which is called the graviton field,

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

This couches general relativity into the framework of Minkowskian quantum field theories. Usually these theories are plagued by infinities. The standard way to cure these so-called UV-divergencies, at least perturbatively, is renormalization. Unfortunately, the quantum field theory for the graviton field is not renormalizable. We also want to mention that this ansatz contradicts our principle of background-independence since $\eta_{\mu\nu}$ fixes a preferred background. But $\eta_{\mu\nu}$ has no intrinsic meaning — only $g_{\mu\nu}$ determines causality.

A more ambitious approach to quantum gravity, which is the most popular one today, is string theory. It is still a perturbative theory and claims the necessity of unifying all interactions to obtain a theory of quantum gravity. It has had remarkable success and, in a certain sense, also predicts that the background metric has to satisfy the Einstein equations. On the other hand, string theory introduces a lot of additional physics like supersymmetry, extra dimensions, an infinite number of universes. So far none of these phenomena have shown up in experiments. Also, to my knowledge, string theory has not yet managed to produce any quantitative predictions which would make the theory falsifiable.

In contrast, loop quantum gravity (LQG), the approach to quantum gravity on which this thesis focuses, is much more conservative. One of its distinguishing features is that it only relies on physically well tested theories, namely general relativity and quantum mechanics. In a sense one tries to find suitable formulations of both theories so that one can straightforwardly quantize general relativity. Therefore an important concept of LQG is background independence. We take Einstein's insight seriously and work without a distinguished background metric from the start. Important successes of LQG are the prediction of discrete spectra of the geometrical operators at the kinematical level, a derivation of the Bekenstein Hawking entropy of black holes (up to a dimensionless constant) and the resolution of the classical big-bang singularity in its application to cosmology.

Part 1 provides an introduction to the subject of loop quantum gravity. The notation and presentation roughly follows the one given by Professor Balasin in his lectures [9] and incorporates a lot of background information and additional material from various sources. My intent was that this part, together with the appendices, could in principle serve as complementary reading material for the lectures. My hope is that in this way one or more students may find easier access to the world of loop quantum gravity and maybe even consider a thesis in the subject.

Part 2 then goes on to discuss loop quantum cosmology (LQC), the quantization of cosmological models with methods of LQG. Current research on this topic focuses mainly on the dynamics. In contrast, our treatment will discuss two approaches to the somewhat neglected kinematical aspects. In chapter 7 we compare the approaches and find that both suffer from certain defects. We then go on to discuss some proposals on how to improve the kinematic structure. The main idea involves implementing invariance under certain diffeomorphisms, which has inspired the name for this thesis.

Part 1

Loop Quantum Gravity

CHAPTER 1

Hamiltonian Formulation of general relativity

As mentioned in the introduction, LQG is an attempt to quantize general relativity by a canonical approach. Thus, we will need a Hamiltonian formulation, which was first developed by Arnowitt, Deser and Misner [2] (a reprint of the original article is available in [3]). A more modern treatment of the topic is given by Wald [42]. It should also be mentioned that the contents of appendices A and B is a prerequisite for all that follows.

We will start with an introductory example, so that hopefully the following results will not come that much as a surprise.

Example 1.1. We consider a classical theory with a finite number of degrees of freedom; for simplicity of presentation, we even restrict ourselves to one coordinate q. The action of the system is given by

$$S[q] = \int L(q, \dot{q}) dt.$$

By a standard Legendre transform

$$H_0(q, p) = p\dot{q} - L(q, \dot{q}), \text{ with } p := \frac{\partial L}{\partial \dot{q}}$$

we pass to the Hamiltonian framework. The trajectory is given by a function $q(t)$ of one independent variable t.

Taking time as an additional observable, this system can equivalently be described by regarding t as a function of an arbitrary parameter τ. Geometrically speaking, we pass from describing the system by the function $q(t)$ to a description in terms of the graph of this function, i.e. a curve in $\mathbb{R}^2$ parameterized by τ. We obtain a system with an extra degree of freedom t. The new action is obtained by using (we denote differentiation w.r.t. τ by a prime and differentiation w.r.t. t by a dot)

$$q' := \frac{dq}{d\tau} = \frac{dq}{dt}\frac{dt}{d\tau} = \dot{q}t',$$

hence

$$S[q] = \int L(q, \dot{q}) dt = \int \underbrace{L(q, \frac{q'}{t'})t'}_{=:L^*(q,t;q',t')} \, d\tau =: S[q, t].$$

Observe that this action is invariant under a reparametrization $\widetilde{\tau} = \widetilde{\tau}(\tau)$.

The momenta in the new description are given by

$$p^* := \frac{\partial L^*}{\partial q'} = \frac{\partial L}{\partial \dot{q}}\frac{1}{t'}t' = p$$

$$p_t := \frac{\partial L^*}{\partial t'} = L + \frac{\partial L}{\partial \dot{q}}\left(-\frac{q'}{t'^2}\right)t' = L - p\dot{q} = -H_0(q,p).$$

We see that p_t is not independent of the other p-s, so that we have the constraint

$$p_t + H_0(q,p) = 0.$$

The Hamiltonian of the "new system" is

$$H = p_t t' + pq' - L^* = p_t t' + p\dot{q}t' - Lt' = t'(p_t + H_0),$$

so it is seen to vanish as a consequence of the constraint. This does not mean that there is no time-evolution. The variable τ is not a time but a mere parameter and our new formalism is invariant under reparametrizations of τ. In the traditional formalism the time t is regarded as a parameter and the trajectory $q(t)$ describes the evolution of the observable q w.r.t. time. What we have done here is to promote t to an additional observable, which is now on equal footing with q. Instead of talking about evolution in a time parameter the new formalism is about *correlations* in the two observables q and t.

We will see later that the Hamiltonian of general relativity is also constrained to vanish. In light of this example we can then interpret this fact as telling us that our theory is given in "already parameterized" form.

1. Nature of the initial value formulation of general relativity

In ordinary field theories on Minkowski space one can specify the field amplitudes and their time derivatives at a given initial time. The field equations then determine the future evolution uniquely.

In general relativity, however, we are immediately facing serious obstacles. The notion of time is obviously observer-dependent, so the question is: What do we mean by "initially"? What is evolution in "time"? Also the choice of phase-space coordinates and velocities is far from obvious, i.e. what is the quantity whose evolution is to be described, and how do we define an initial time derivative of this quantity? Finally we have to take general covariance into account — there is no distinguished time variable, so a reparametrization-independent formulation as described in the initial example will be appropriate. Since general relativity is invariant under spacetime diffeomorphisms, we can always make a coordinate transformation at any later time. This will change the metric field, but not yield any observable physical change. So we will have to find a way to isolate the "true dynamical degrees of freedom" from those merely describing the choice of specific coordinates.

The first intuition comes from our discussion of hypersurfaces in appendix A. We saw in formula (A.7) that the extrinsic curvature is related to the derivative of the induced metric. This raises the idea that h_{ab} might play the role of coordinates and K_{ab} is related to velocities. We will now make this intuition precise.

We assume that our spacetime (M, g_{ab}) is globally hyperbolic. Roughly speaking this means that there exists a spacelike hypersurface Σ (called a Cauchy surface) such that the entire history of the universe is determined by fixing the initial conditions on Σ. Since predictability is an essential feature of an initial value formulation, our assumption does not seem to be very restrictive. It then follows that spacetime foliates into Cauchy surfaces Σ_t, parameterized by a global time function $t\colon M \to \mathbb{R}$ (t is strictly increasing along every future directed timelike curve). An outline of how to prove this assertion can by found in theorem 8.3.14 of [42], where one can also find precise definitions of Cauchy surfaces and global hyperbolicity.

To deal with the arbitrariness of choosing coordinates we introduce a global "time-flow" vector-field t^a, subject only to the condition

$$t^a \nabla_a t = 1. \tag{1.1}$$

t^a represents the flow of time and the condition ensures that it is nowhere tangential to Σ_t (so that time "does not stop")[1]. We can then identify the surfaces Σ_0 and Σ_t along the integral curves of t^a. Dynamics then consist in describing the evolution in parameter time t of quantities on an abstract three-manifold $\Sigma := \Sigma_0$. Since in general relativity the quantity of interest is the metric, the natural object to consider is the induced metric h_{ab} (A.2).

The arbitrariness in the choice of the vector field t^a is parameterized by the lapse-function N and shift-vector N^a, defined as its normal and tangential components, respectively (see also Figure 1),

$$t^a = N n^a + N^a. \tag{1.2}$$

The lapse-function N determines the change of proper time of an observer with 4-velocity n^a versus coordinate time t, whereas the shift-vector N^a measures how the coordinate system changes when we go from Σ_t to $\Sigma_{t'}$. Note that since N is a scalar object and N^a is tangential, both are intrinsic to the manifold Σ.

Given this foliation, it is convenient to work with coordinates that are adapted to this special geometric situation. The original choice with arbitrary coordinates x^i on Σ_t, identified along the integral curves of t^a with parameter t has the drawback of ∂_i^a not being orthogonal to ∂_t^a. We can easily correct this by using instead ∂_i^a and $\partial_t^a - N^i \partial_i^a$ as coordinate vector fields. Since the latter equals $t^a - N^a = N n^a$, orthogonality is obvious. The change to orthogonal coordinates, along with the corresponding change of dual bases, is illustrated in the following diagram,

$$\begin{array}{c:c} \partial_t^a & dt \\ \hdashline \partial_i^a & dx^i \end{array} \quad \rightsquigarrow \quad \begin{array}{c:c} \partial_t^a - N^i \partial_i^a & dt \\ \hdashline \partial_i^a & dx^i + N^i dt \end{array} \quad .$$

[1]Note that t^a is allowed to be spacelike since it merely prescribes the change of the coordinate system in time.

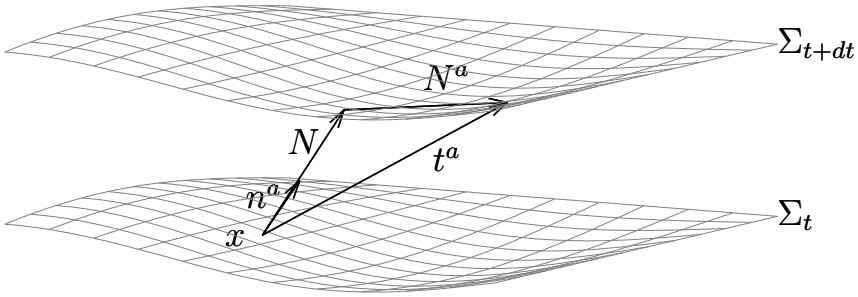

FIGURE 1. The definition of lapse and shift. t^a connects points of the same spatial coordinates x and is decomposed into the tangential part N^a and N times the unit normal.

We can use our new coordinates to find a proper definition of velocities. Our first intuition would be to use $\mathcal{L}_t h_{ab}$, but this object will in general not be tangential, so we have to be more careful. The obvious way to correct this is to project both indices with $h_a{}^b$. This is indeed seen to be appealing by writing the induced metric in above coordinates (remember that the normal component of tangential objects is zero),

$$h_{ab} = h_{ij}(dx^i + N^i dt)_a (dx^j + N^j dt)_b. \tag{1.3}$$

We expand this expression and act on it with $\mathcal{L}_t$. Using the coordinate covariant derivative ∂_a and a coordinate vector-field ∂_i^a in the formula $\mathcal{L}_v t_{ab} = v^c \nabla_c t_{ab} + t_{cb}\nabla_a v^c + t_{ac}\nabla_b v^c$, we see that $\mathcal{L}_{\partial_i} t_{ab} = \partial_i t_{ab}$, since $\partial_i^a \partial_a = \partial_i$ and $\partial_a \partial_i^b = 0$. Hence,

$$\mathcal{L}_t h_{ab} = \partial_t h_{ij} dx_a^i dx_b^j + 2\partial_t \left(h_{ij} N^i\right) dt_a dx_b^j + \partial_t \left(h_{ij} N^i N^j\right) dt_{ab}^2.$$

The second and third term vanish when we act with the projection operator, because dt_a is proportional to n_a, see equation (A.1). Thus, we arrive at

$$h_a{}^c h_b{}^d \mathcal{L}_t h_{cd} = \dot{h}_{ij} dx_a^i dx_b^j.$$

This expression tells us that our first intuition for velocities is indeed correct, since the right-hand side certainly is the appropriate definition when the intrinsic metric is viewed as a three-dimensional object. Concluding, velocities are defined via

$$\dot{h}_{ab} := h_a{}^c h_b{}^d \mathcal{L}_t h_{cd}.$$

We are now ready to summarize our findings and give an interpretation of the formalism. We started with a globally hyperbolic spacetime (M, g_{ab}), which foliates into hypersurfaces Σ_t, each of which carries the induced metric $h_{ab}(t)$. The arbitrary time-flow vector-field allowed us to identify all Σ_t-s with $\Sigma := \Sigma_0$. After deriving the EOM satisfied by $h_{ab}(t)$, we can thus forget about the original spacetime and arrive at the following initial value formulation: GR describes the evolution of a Riemannian metric h_{ab} on a fixed manifold Σ.[2] Given lapse-function and shift-vector, as well as (suitable[3]) initial data $h_{ab}(0)$ and $\dot{h}_{ab}(0)$ (or equivalently K_{ab}, see equation(1.8b)), we can use the EOM (1.22) to calculate $h_{ab}(t)$. The chosen lapse and shift tell us how to glue the different slices together to obtain our original spacetime, and the spacetime metric can be calculated via ((A.2) with (1.3) and (1.5)),

$$g_{ab} = -N^2 dt_{ab}^2 + h_{ij}(dx^i + N^i dt)_a (dx^j + N^j dt)_b. \tag{1.4}$$

We will see in section 4, that if we start out with different lapse and shift but with the same initial data, we obtain a diffeomorphic spacetime.

2. Canonical form of the action

To derive the EOM in canonical form we start by rewriting the action (B.12) as a functional of induced metric and extrinsic curvature, as well as lapse and shift.

The expression for the curvature scalar is derived in appendix A, the final result is (A.11). For the volume element, we contract equation (1.2) with n_a to obtain $N = -n_a t^a$. Inserting $n = \alpha dt$ (see (A.1)) and using (1.1) yields $N = -\alpha$, giving

$$n_a = -N \nabla_a t. \tag{1.5}$$

This equation allows us to write the volume element as

$$\omega_g = \sigma N \omega_h \wedge dt, \tag{1.6}$$

by using the expression (A.14).

As already mentioned, the extrinsic curvature appearing in (A.11) is directly related to $\dot{h}_{ab}$.

$$\begin{aligned} K_{ab} &= \frac{1}{2}\mathcal{L}_n h_{ab} = \frac{1}{2}\mathcal{L}_{\frac{1}{N}(t^c - N^c)} h_{ab} \\ &\overset{(*)}{=} \frac{1}{2N}\left(\mathcal{L}_{t^c} - \mathcal{L}_{N^c}\right) h_{ab} \\ &= \frac{1}{2N}\left(h_a{}^c h_b{}^d \mathcal{L}_{t^c} h_{cd} - h_a{}^c h_b{}^d \mathcal{L}_{N^c} h_{cd}\right), \end{aligned} \tag{1.7}$$

[2]Since the full metric g_{ab} is completely specified by h_{ab}, N and N^a, we should also view lapse and shift as dynamical variables. However, we will see in section 3, that their evolution is arbitrary and they should rather be regarded as describing the foliation into space and time.

[3]The initial data have to satisfy the constraint equations (1.13) and (1.11), but apart from that they can be arbitrary symmetric tensor-fields. In particular, K_{ab} is a priori not interpreted as extrinsic curvature, since Σ is not embedded into a larger spacetime.

where $n^a = \frac{1}{N}\left(t^a - N^a\right)$ was used in the second equality and we used the fact that K_{ab} is tangential to insert the projectors. In the equality marked with $(*)$ we pulled out $\frac{1}{N}$ from the Lie derivative, although this factor is not constant. We can justify this as follows,

$$\begin{aligned}\mathcal{L}_{\frac{1}{N}(t^c-N^c)}h_{ab} &= \frac{1}{N}\left(\left(t^c - N^c\right)\nabla_c\right)h_{ab} \\ &\quad + \nabla_a\left(\frac{1}{N}\left(t^c - N^c\right)\right)h_{cb} + \nabla_b\left(\frac{1}{N}\left(t^c - N^c\right)\right)h_{ac}.\end{aligned}$$

We observe that when ∇_a acts on $\frac{1}{N}$ in the second term we are left with $\left(t^c - N^c\right)h_{cd} = n^c h_{cd} = 0$ and similarly for the third term. We can thus pull out $\frac{1}{N}$ from under the derivative in these terms and are left with $\frac{1}{N}\mathcal{L}_{(t^c-N^c)}h_{ab}$.

The first term on the right-hand side of equation (1.7) can already be identified as $\dot{h}_{ab}$, whereas the second term has to be further modified.

$$\begin{aligned}h_a{}^c h_b{}^d \mathcal{L}_{N^c} h_{cd} &= h_a{}^c h_b{}^d \left(N^e \nabla_e\right) h_{cd} + h_a{}^c h_b{}^d \left(\nabla_c N^e\right) h_{ed} + h_a{}^c h_b{}^d \left(\nabla_d N^e\right) h_{ce} \\ &= D_a N_b + D_b N_a,\end{aligned}$$

where we used $h_{ab} = g_{ab} + n_a n_b$ in all three terms, together with metric compatibility of ∇_a and $h_a{}^c n_c = 0$ to see that the first term vanishes and the others can be written as indicated.

We thus have the following relation between extrinsic curvature and the time derivative of the intrinsic metric,

$$K_{ab} = \frac{1}{2N}\left(\dot{h}_{ab} - D_a N_b - D_b N_a\right).$$

Putting our results together, we can rewrite the action of GR as follows.

$$\begin{aligned}S_{\text{EHH}} &= \int_M R\omega_g + 2\sigma\int_{\partial M} K\omega_h \\ &= \int_M \left({}^3R - K^2 + K_{ab}K^{ab} - 2\nabla_a\left((n\nabla)n^a\right) + 2\nabla_a\left(n^a K\right)\right)\omega_g + 2\sigma\int_{\partial M} K\omega_h \\ &= \int_M \left({}^3R - K^2 + K_{ab}K^{ab}\right)\sigma N\omega_h \wedge dt - 2\sigma\int_{\partial M} n_a\left((n\nabla)n^a\right)\omega_h \\ &\quad + 2\sigma\int_{\partial M} n_a n^a K\omega_h + 2\sigma\int_{\partial M} K\omega_h,\end{aligned}$$

by Gauß's law (A.17). The second term vanishes due to $n_a n^a = -1$ and the Leibniz rule; the third and fourth term cancel. We arrive at the canonical form of the action,

$$S_{\text{EHH}} = \int_{\mathbb{R}} dt \underbrace{\int_\Sigma \left({}^3R - K^2 + K_{ab}K^{ab}\right) N\omega_h}_{=:L_{\text{ADM}}(h_{ab},\dot{h}_{ab})} \tag{1.8a}$$

$$\text{with} \quad K_{ab} = \frac{1}{2N}\left(\dot{h}_{ab} - D_a N_b - D_b N_a\right) \tag{1.8b}$$

3. Legendre transform

From this point on the reader should have some knowledge of symplectic manifolds. It is therefore a good idea to consult appendix C before continuing.

Our first observation is that L_{ADM} does not depend on $\dot{N}$ and $\dot{N}^a$.[4] The Euler-Lagrange equations for lapse and shift thus yield $\frac{\delta L_{\mathrm{ADM}}}{\delta N} = 0$ and $\frac{\delta L_{\mathrm{ADM}}}{\delta N^a} = 0$, that is they tell us that our theory is independent of the chosen foliation into space and time. We examine those equations in detail to see what they imply.

For the variation with respect to N, we need

$$\delta_N K_{ab} = -\frac{1}{2N^2}\delta N \left(\dot{h}_{ab} - D_a N_b - D_b N_a\right) = -\frac{\delta N}{N} K_{ab}. \tag{1.9}$$

The variation of the Lagrangian therefore is

$$\begin{aligned}
\delta_N L_{\mathrm{ADM}} &= \int \left({}^3R - K^2 + K_{ab}K^{ab}\right)\delta N \omega_h + 2\int \left(-\delta_N K_{ab} h^{ab} K + \delta_N K_{ab} K^{ab}\right) N\omega_h \\
&= \int \left(\left({}^3R - K^2 + K_{ab}K^{ab}\right) + 2\left(K^2 - K_{ab}K^{ab}\right)\right)\delta N \omega_h \\
&= \int \left({}^3R + K^2 - K_{ab}K^{ab}\right)\delta N \omega_h,
\end{aligned} \tag{1.10}$$

from which we can read off $\frac{\delta L_{\mathrm{ADM}}}{\delta N}$ and write the Euler-Lagrange equation as

$$\widetilde{H} := \left(K_{ab}K^{ab} - K^2 - {}^3R\right)\omega_h = 0. \tag{1.11}$$

Since this equation imposes a constraint on the canonical variables, it is known as *Hamiltonian constraint*. The sign in the definition of $\widetilde{H}$ is chosen so that we arrive at (1.21) and consequently have $\widetilde{H} = \frac{\delta H_{\mathrm{ADM}}}{\delta N}$.

For the variation with respect to N^a we use

$$\delta_{N^c} K_{ab} = \frac{1}{2N}\left(-D_a \delta N_b - D_b \delta N_a\right).$$

Since we demand $\delta_{N^a} L_{\mathrm{ADM}} = 0$ for *all* N^a, we can in particular consider shift-vectors of compact support, which means we can ignore boundary terms arising from partial

[4]When we try to define conjugate momenta for lapse and shift via $\pi := \frac{\delta \mathcal{L}}{\delta \dot{N}}$, resp. $\pi_a := \frac{\delta \mathcal{L}}{\delta \dot{N}^a}$, we are immediately facing the problem, that we cannot solve the velocities $\dot{N}$ and $\dot{N}^a$ for momenta, because obviously $\pi = \pi_a = 0$. Strictly speaking, we are dealing with a theory with constraints and should use Dirac's algorithm. However, the constraints have a very easy structure in our case, so we can make a convenient shortcut here and appeal to physical intuition. The reader interested in Dirac's ideas is referred to [24], a more detailed exposition is found e.g. in [26].

integrations.[5] Using the symmetry of K^{ab} and h^{ab} we get

$$\begin{aligned}\delta_{N^c} L_{\text{ADM}} &= 2\int \left(-\delta_{N^c} K_{ab} h^{ab} K + \delta_{N^c} K_{ab} K^{ab}\right) N\omega_h \\ &= 2\int \left(\frac{1}{2N} h^{ab} K D_a \delta N_b - \frac{1}{2N} K^{ab} D_a \delta N_b\right) N\omega_h \\ &= \int \left(-D_b K + D_a K^a{}_b\right) \delta N^b \omega_h,\end{aligned} \tag{1.12}$$

by lowering the index b on K^{ab} and h^{ab} in exchange for raising the index in δN_b to go to the last line. This time we get the *vector-* or *diffeomorphism constraint*

$$\widetilde{V}_b := 2\left(D_b K - D_a K^a{}_b\right)\omega_h = 0. \tag{1.13}$$

Since the Euler-Lagrange equations for lapse and shift merely yielded constraints, the evolution of N and N^a is completely arbitrary. In other words, lapse and shift are not dynamical after all, they only parameterize the foliation into space and time. It is therefore natural to perform the Legendre transform with respect to the remaining variable h_{ab}. In other words, our phase space will be the cotangent bundle $T^*\mathcal{C}$, where $\mathcal{C}$ is the space of all three-dimensional Riemannian metrics on Σ. Since this space is infinite-dimensional, a precise definition is hard to give. We will proceed the same way most physicist do and assume that the cotangent bundle has coordinates h_{ab} in the base and let $\widetilde{\pi}^{ab}$ denote the fiber coordinates.

The conjugate momentum is thus defined by $\widetilde{\pi}^{ab} := \frac{\delta L_{\text{ADM}}}{\delta \dot{h}_{ab}}$.[6] The functional derivative is calculated via the variation

$$\begin{aligned}\delta_{\dot{h}} L_{\text{ADM}} &= \int \left(2K^{ab}\frac{1}{2N}\delta\dot{h}_{ab} - 2Kh^{ab}\frac{1}{2N}\delta\dot{h}_{ab}\right) N\omega_h \\ &= \int \left(K^{ab} - h^{ab}K\right)\delta\dot{h}_{ab}\omega_h =: \int \widetilde{\pi}^{ab}\delta\dot{h}_{ab},\end{aligned} \tag{1.14}$$

where we used $\delta_{\dot{h}_{ab}} K_{ab} = \frac{1}{2N}\delta\dot{h}_{ab}$. Observe that in order for the last integral to make sense without using any background structure (i.e. without an arbitrary background volume element), we have to define the conjugate momentum as a tensor *density*. Let us again highlight this important definition,

$$\widetilde{\pi}^{ab} := \frac{\delta L_{\text{ADM}}}{\delta \dot{h}_{ab}} = \left(K^{ab} - h^{ab}K\right)\omega_h. \tag{1.15}$$

[5]The geometrical meaning of a shift-vector of compact support is, that the coordinate system only changes on a compactly supported part of Σ in going from Σ_t to Σ_{t+dt}. Note that we could not take a compactly supported lapse-function, since this would imply that Σ_t equals Σ_{t+dt} outside the compact part, which is not a proper foliation!

[6]We will work with the following definition of the functional derivative $\frac{\delta f}{\delta \phi}$. A functional f on a manifold of fields ϕ is said to be functionally differentiable, if there exists a tensor density $\frac{\delta f}{\delta \phi}$ dual to ϕ (i.e. with index structure such that all indices in the following integral are contracted), such that $\delta f = \int \frac{\delta f}{\delta \phi}\delta\phi$. Here we also allow $\frac{\delta f}{\delta \phi}$ to be distributional, as will be the case for example for $\frac{\delta h_{ab}}{\delta h_{cd}}$.

We can invert this definition and express K_{ab} in terms of $\widetilde{\pi}^{ab}$ by contracting with h_{ab},

$$\widetilde{\pi} = (K - 3K)\,\omega_h = -2K\omega_h. \tag{1.16}$$

Reinserting the contracted equation into the definition yields

$$K^{ab} = \left(\widetilde{\pi}^{ab} - \frac{1}{2}\widetilde{\pi}h^{ab}\right)\omega_h^{-1} \quad \text{and} \tag{1.17}$$

$$K = -\frac{1}{2}\widetilde{\pi}\omega_h^{-1}. \tag{1.18}$$

This allows for expressing the constraints in canonical variables. For the Hamiltonian constraint we calculate

$$K_{ab}K^{ab} - K^2 - {}^3R = \left(\widetilde{\pi}^{ab}\widetilde{\pi}_{ab} - \widetilde{\pi}^2 + \frac{3}{4}\widetilde{\pi}^2 - \frac{1}{4}\widetilde{\pi}^2\right)\omega_h^{-2} - {}^3R.$$

We multiply by ω_h to obtain the Hamiltonian constraint

$$\widetilde{H} = \left(\widetilde{\pi}^{ab}\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}^2\right)\omega_h^{-1} - {}^3R\omega_h. \tag{1.19}$$

For the vector constraint

$$(D_aK^a{}_b - D_bK)\,\omega_h = D_a\widetilde{\pi}^a{}_b - \frac{1}{2}D_a\widetilde{\pi}h^a{}_b + D_b\frac{1}{2}\widetilde{\pi} = D_a\widetilde{\pi}^a{}_b,$$

and therefore

$$\widetilde{V}_b = -2D_a\widetilde{\pi}^a{}_b. \tag{1.20}$$

We can now perform the Legendre transform and obtain for the ADM-Hamiltonian

$$\begin{aligned}
H_{\text{ADM}} &= \int \widetilde{\pi}^{ab}\dot{h}_{ab} - L_{\text{ADM}} = \int \widetilde{\pi}^{ab}\left(2NK_{ab} + 2D_aN_b\right) - L_{\text{ADM}} \\
&= \int \widetilde{\pi}^{ab}\left(2N\left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}h_{ab}\right)\omega_h^{-1} + 2D_aN_b\right) \\
&- \int \left(\left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}h_{ab}\right)\left(\widetilde{\pi}^{ab} - \frac{1}{2}\widetilde{\pi}h^{ab}\right)\omega_h^{-2} - \frac{1}{4}\widetilde{\pi}^2\omega_h^{-2} + {}^3R\right)N\omega_h \\
&= \int \left[\left(2\widetilde{\pi}^{ab}\widetilde{\pi}_{ab} - \widetilde{\pi}^2 - \widetilde{\pi}_{ab}\widetilde{\pi}^{ab} + \widetilde{\pi}^2 - \frac{3}{4}\widetilde{\pi}^2 + \frac{1}{4}\widetilde{\pi}^2\right)N\omega_h^{-1} - {}^3RN\omega_h\right] - \int 2N^bD_a\widetilde{\pi}^a{}_b \\
&= \int N\left[\left(\widetilde{\pi}^{ab}\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}^2\right)\omega_h^{-1} - {}^3R\omega_h\right] - \int 2N^bD_a\widetilde{\pi}^a{}_b \\
&= \int \widetilde{H}N + \int \widetilde{V}_bN^b.
\end{aligned}$$

using (1.8b) and the obvious symmetry of $\widetilde{\pi}^{ab}$ in the first line, then inserting equations (1.17) and (1.18) and the Lagrangian (1.8a). We continued by factoring out (the numerator in the $\frac{3}{4}$ comes from $h_{ab}h^{ab} = 3$) and performing a partial integration, dropping again the boundary terms by restricting ourselves to a shift-vector of compact support.

One often uses the following abbreviated notation for the ADM-Hamiltonian,

$$H_{\text{ADM}} = H(N) + V(N^c). \tag{1.21}$$

4. Motions generated by the constraints

The EOM for the canonical variables are (cf. (C.9))

$$\dot{h}_{ab} = \{h_{ab}, H_{\text{ADM}}\} \tag{1.22a}$$

$$\dot{\widetilde{\pi}}^{ab} = \left\{\widetilde{\pi}^{ab}, H_{\text{ADM}}\right\}. \tag{1.22b}$$

In light of equation (1.21) and the additivity of the Poisson-bracket, we can split the evolution into a part generated by the diffeomorphism constraint and one generated by the Hamiltonian constraint.

We first examine the contribution from the diffeomorphism constraint. The Poisson-bracket is computed by equation (C.8), hence we need the variation of $V(N^c)$. We will first simplify the constraint, assuming compact Σ with $\partial\Sigma = 0$ to be able to drop boundary terms in the partial integration.

$$\begin{aligned} V(N^c) &= -2\int N^c D_a \widetilde{\pi}^a{}_c = 2\int D_a N_b \widetilde{\pi}^{ab} \\ &= \int \left(D_a N_b + D_b N_a\right)\widetilde{\pi}^{ab} = \int \left(\mathcal{L}_{N^c} h_{ab}\right)\widetilde{\pi}^{ab}, \end{aligned} \tag{1.23}$$

using metric compatibility to change index positions and then exploiting the symmetry of $\widetilde{\pi}^{ab}$ to obtain the formula for the Lie derivative. Now we can easily read off the needed variations (after a partial integration for δ_h[7])

$$\delta_h V(N^c) = -\int \mathcal{L}_{N^c}\widetilde{\pi}^{ab}\delta h_{ab} \tag{1.24a}$$

$$\delta_{\widetilde{\pi}} V(N^c) = \int \mathcal{L}_{N^c} h_{ab}\delta\widetilde{\pi}^{ab} \tag{1.24b}$$

and find the desired expression for the Poisson-brackets

$$\{h_{ab}, V(N^c)\} = \int \frac{\delta h_{ab}}{\delta h_{cd}}\frac{\delta V(N^c)}{\delta\widetilde{\pi}^{cd}} = \mathcal{L}_{N^c} h_{ab} \tag{1.25a}$$

$$\left\{\widetilde{\pi}^{ab}, V(N^c)\right\} = -\int \frac{\delta\widetilde{\pi}^{ab}}{\delta\widetilde{\pi}^{cd}}\frac{\delta V(N^c)}{\delta h_{cd}} = \mathcal{L}_{N^c}\widetilde{\pi}^{ab}. \tag{1.25b}$$

For the motions generated by the Hamiltonian constraint we use expression (1.19) to find

$$\delta_{\widetilde{\pi}} H(N) = 2\int \left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi} h_{ab}\right) N\omega_h^{-1}\delta\widetilde{\pi}^{ab},$$

[7]Partial integration for Lie derivatives can be obtained as usual by employing the Leibniz rule. The term $\int \mathcal{L}_{N^c}\left(h_{ab}\widetilde{\pi}^{ab}\right)$ can be dealt with by using $\mathcal{L}_v\omega = v \lrcorner d\omega + d(v \lrcorner \omega)$ for arbitrary forms ω. In our case the Lie derivative acts on a scalar density, which is a form of highest degree. The first term thus vanishes, the second can be written as a boundary term by applying Stokes's theorem.

from which we infer by equations (1.15) and (1.16)

$$\begin{aligned}\{h_{ab}, H(N)\} &= 2\left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}h_{ab}\right) N\omega_h^{-1} \\ &= 2\left(K_{ab} - Kh_{ab} + Kh_{ab}\right)\omega_h N\omega_h^{-1} \\ &= 2NK_{ab} = 2N\frac{1}{2}\mathcal{L}_{n^c}h_{ab} \\ &= \mathcal{L}_{Nn^c}h_{ab}.\end{aligned} \tag{1.26}$$

For the reason why we could put the (non-constant) lapse N into the Lie-derivative, see the corresponding argument in equation (1.7). The analog result for the momentum holds only when the vacuum Einstein equations $G_{ab} = 0$ are satisfied and is much more complicated to derive (see [40] for a sketch of how to show this result).

We are now ready to give a geometrical interpretation of the motions generated by the constraints. From (1.25) we see that $V(N^c)$ generates infinitesimal diffeomorphisms of Σ along the integral curves of the shift vector N^c. This was to be expected from the fact that the shift vector is related to the choice of the coordinate system on Σ.

Correspondingly, equation (1.26) and the analog result for $\widetilde{\pi}^{ab}$ show that $H(N)$ generates diffeomorphisms orthogonal to Σ_t. This is also the origin of the name *Hamiltonian* constraint, since evolution orthogonal to Σ_t is related to evolution in "time". It should however be emphasized that for this interpretation we had to bring in some aspects of the four-dimensional spacetime. Note also that the interpretation for the Hamiltonian constraint only holds when the vacuum Einstein equations are satisfied.

As advertised above, now we also see that starting out with different lapse and shift yields a diffeomorphic spacetime. We denote the Hamiltonian vector field generating evolution by $X_{H_{\mathrm{ADM}}}$ and have

$$\begin{aligned}X_{H(N)+V(N^c)}h_{ab} - X_{H(M)+V(M^c)}h_{ab} &= X_{H(N-M)+V(N^c-M^c)}h_{ab} \\ &= \mathcal{L}_{(N-M)n^c+(N^c-M^c)}h_{ab}.\end{aligned}$$

5. Hypersurface deformation algebra

An important consistency requirement is that the constraints be preserved under evolution in "parameter-time".[8] After all, if the constraints were only satisfied on a particular spatial hypersurface Σ_t, this would mark a preferred initial time — certainly an undesirable feature we want to avoid.

Since the evolution of functions on phase-space is governed by equations (1.22), we have to show

$$\dot{H}(N) = \{H(N), H_{\mathrm{ADM}}\} \approx 0 \tag{1.27}$$

$$\dot{V}(N^c) = \{V(N^c), H_{\mathrm{ADM}}\} \approx 0, \tag{1.28}$$

[8]In Dirac's terminology, we have to show that the constraints are first class rather than second class.

where we have introduced the notation $\approx$ to denote "equal, provided that the constraints hold". In light of equation (1.21), it is sufficient to show that the constraint algebra closes, i.e. a Poisson bracket of two constraints yields another constraint. We will now prove that the constraints fulfill the following algebra, known as *Dirac algebra*, sometimes also called *hypersurface deformation algebra*,

$$\left\{V(N^c), V(M^d)\right\} = V\left(\mathcal{L}_{N^c} M^d\right) \tag{1.29a}$$

$$\{V(M^c), H(N)\} = H\left(\mathcal{L}_{M^c} N\right) \tag{1.29b}$$

$$\{H(N), H(M)\} = V\left(\left(N D_a M - M D_a N\right) h^{ab}\right) \tag{1.29c}$$

For the first Poisson-bracket, we can use equations (C.8) and read off the needed variational derivatives from expressions (1.24). By a partial integration we obtain

$$\begin{aligned}\left\{V(N^c), V(M^d)\right\} &= \int -\mathcal{L}_{N^c}\widetilde{\pi}^{ab}\mathcal{L}_{M^d} h_{ab} + \mathcal{L}_{M^d}\widetilde{\pi}^{ab}\mathcal{L}_{N^c} h_{ab} \\ &= \int \left[\mathcal{L}_{N^c}, \mathcal{L}_{M^d}\right] h_{ab}\widetilde{\pi}^{ab} = \int \mathcal{L}_{\left[N^c, M^d\right]} h_{ab}\widetilde{\pi}^{ab} \\ &= V\left(\left[N^c, M^d\right]\right).\end{aligned}$$

The calculation of the second bracket can be simplified tremendously by the following observation (we denote $H_N := H(N)$ and $V_M := V(M^c)$ to save space).

$$\begin{aligned}\{V(M^c), H(N)\} &= \int \left(\frac{\delta V_M}{\delta h_{ab}}\frac{\delta H_N}{\delta \widetilde{\pi}^{ab}} - \frac{\delta V_M}{\delta \widetilde{\pi}^{ab}}\frac{\delta H_N}{\delta h_{ab}}\right) \\ &= \delta_{\widetilde{\pi}} H_N\big|_{\delta\widetilde{\pi}^{ab} = \frac{\delta V_M}{\delta h_{ab}}} - \delta_h H_N\big|_{\delta h_{ab} = \frac{\delta V_M}{\delta \widetilde{\pi}^{ab}}},\end{aligned} \tag{1.30}$$

where the variations in the second line can be interpreted as derivations into an unspecified direction (very much like a gradient), and the functional derivatives of the diffeomorphism constraint specify the direction.

Since in our case the functional derivative of the diffeomorphism constraint takes the form of a Lie derivative (equations (1.24)), the last expression can be rewritten as a Lie derivative. The crucial observation is that since the variation acts as a derivation, we can apply the chain rule. For example,

$$\delta_{\widetilde{\pi}} H_N = \delta_{\widetilde{\pi}} \int N\left(\left(\widetilde{\pi}^{ab}\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}^2\right)\omega_h^{-1} - {}^3R\omega_h\right) = \int 2N\left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi} h_{ab}\right)\omega_h^{-1}\delta\widetilde{\pi}^{ab}. \tag{1.31}$$

Now when we calculate the Lie derivative $\mathcal{L}_{M^c} H_N$, we *also* use the chain rule, only this time the inner derivative is not $\delta\widetilde{\pi}^{ab}$, but $\mathcal{L}_{M^c}\widetilde{\pi}^{ab}$, and correspondingly for the derivative w.r.t. h_{ab}. We only have to be careful because in contrast to the variation, the chain rule for the Lie derivative also contains a term where $\mathcal{L}_{M^c}$ acts on N. But we can compensate this by expanding $H_N = \int \widetilde{H} N$ and letting the Lie derivative act only on $\widetilde{H}$. This trick allows us to compute the Poisson bracket without ever calculating δH_N. From (1.30) we

continue with

$$
\begin{aligned}
&-\left(\delta_{\widetilde{\pi}} H_N|_{\delta\widetilde{\pi}^{ab}=\mathcal{L}_{M^c}\widetilde{\pi}^{ab}} + \delta_h H_N|_{\delta h_{ab}=\mathcal{L}_{M^c}h_{ab}}\right)\\
&= -\int N\mathcal{L}_{M^c}\widetilde{H} = \int \mathcal{L}_{M^c} N\widetilde{H}\\
&= H\left(\mathcal{L}_{M^c}N\right),
\end{aligned}
$$

again by a partial integration in the second line.

The calculation of the Poisson bracket between the two Hamiltonian constraints is the most laborious one. $\delta_{\widetilde{\pi}} H_N$ was already calculated in (1.31). The variation w.r.t. h on the other hand is rather complicated (especially due to the appearance of R and ω_h^{-1}), so we want to avoid its calculation as much as possible. Since

$$
\{H_N, H_M\} = \int \left(\frac{\delta H_N}{\delta h_{ab}}\frac{\delta H_M}{\delta\widetilde{\pi}^{ab}} - \frac{\delta H_N}{\delta\widetilde{\pi}^{ab}}\frac{\delta H_M}{\delta h_{ab}}\right) \tag{1.32}
$$

is antisymmetric in N and M, we will start by calculating the first term and only in the end antisymmetrize. As in the derivation of the previous Poisson bracket, we will write this first term as a directional derivative,

$$
\delta_h H_N|_{\delta h_{ab}=\frac{\delta H_M}{\delta\widetilde{\pi}^{ab}}}\,. \tag{1.33}
$$

The first useful observation is that $\delta_h H_N$ is linear in N, and $\frac{\delta H_M}{\delta\widetilde{\pi}^{ab}}$ is linear in M. Consequently, all terms of $\delta_h H_N$ in which there is no action of D_a on δh_{ab} will be proportional to NM after the substitution $\delta h_{ab} = \frac{\delta H_M}{\delta\widetilde{\pi}^{ab}}$. Since these terms are obviously symmetric, we can neglect them. Taking another look at equation (1.19), we see that the first term can be neglected, since ω_h^{-1} is an algebraic expression of h_{ab} (i.e. it contains no derivatives of h_{ab}). We are left with the task of computing $\delta_h(N^3 R\omega_h)$, which we already did in equation (B.8). Inserting the factor N, which we did not have back then, we get

$$
\delta_h\left(N^3 R\omega_h\right) = -\int NG^{ab}\delta h_{ab}\omega_h + \int Nh^{ab}\delta R_{ab}\omega_h.
$$

The first term again yields a contribution symmetric in N and M and will be discarded. Observe that in contrast to the derivation of the Einstein-Hilbert action we can not cast the second term into a surface term due to the appearance of the extra factor N. Expression (B.9) is just as well valid in three dimensions. We therefore have

$$
\delta_h H_N = -\int N\left(D^a D^b \delta h_{ab} - D^2\delta h\right)\omega_h,
$$

apart from terms that drop out after antisymmetrization anyway. Inserting the direction $\delta h_{ab} = \frac{\delta H_M}{\delta\widetilde{\pi}^{ab}}$ from equation (1.31) and using $h_{ab}h^{ab} = 3$ for the contracted version we are left with

$$
\delta_h H_N|_{\delta h_{ab}=\frac{\delta H_M}{\delta\widetilde{\pi}^{ab}}} = -\int N\left(D^a D^b\left(2M\omega_h^{-1}\left(\widetilde{\pi}_{ab} - \frac{1}{2}\widetilde{\pi}h_{ab}\right)\right) + D^2\left(2M\omega_h^{-1}\frac{1}{2}\widetilde{\pi}\right)\right)\omega_h. \tag{1.34}
$$

Due to metric compatibility of D_a we can pull out ω_h^{-1} from under the derivative and cancel with ω_h. The two terms containing the contracted $\widetilde{\pi}$ cancel and we arrive at

$$\begin{aligned}\int -2ND^aD^b\left(M\widetilde{\pi}_{ab}\right) &= -2\int ND^a\left(D^bM\widetilde{\pi}_{ab}+MD^b\widetilde{\pi}_{ab}\right)\\ &= -2\int\left(ND^aD^bM\widetilde{\pi}_{ab}+2ND^bMD^a\widetilde{\pi}_{ab}+NMD^aD^b\widetilde{\pi}_{ab}\right)\end{aligned}$$

by applying the Leibniz rule twice and remembering the symmetry of $\widetilde{\pi}^{ab}$ for the second application. The last term is symmetric in M and N and will be discarded. In the first term we will perform a partial integration for D^a. There will be one symmetric term proportional to $D^aND^bM\widetilde{\pi}_{ab}$ that can also be dropped. We are left with

$$-2\int\left(-ND^bMD^a\widetilde{\pi}_{ab}+2ND^bMD^a\widetilde{\pi}_{ab}\right)=-2\int ND^bMD^a\widetilde{\pi}_{ab}.$$

Finally we can antisymmetrize this expression to obtain

$$\{H(N),H(M)\}=-2\int\left(ND^bM-MD^bN\right)D_a\widetilde{\pi}^a{}_b=V\left(\left(ND_aM-MD_aN\right)h^{ab}\right),$$

where metric compatibility of D_a was used to exchange index positions and the diffeomorphism constraint was used in the form (1.23).

Let us now see how we can interpret the Dirac algebra. We saw in the previous section that the diffeomorphism constraint generates spatial diffeomorphisms, in other words evolution with $V(N^c)$ yields a new metric $\widetilde{h}_{ab}=X_{V(N^c)}h_{ab}$ that is diffeomorphic to h_{ab}. According to equation (C.4) the commutator of two evolutions $X_{V(N^c)}$ and $X_{V(M^d)}$ is an evolution with $X_{\left\{V(N^c),V(M^d)\right\}}$, which according to the first equation of the Dirac algebra (1.29a) equals another spatial diffeomorphism. It should now be obvious that also the commutator of two evolutions with a diffeomorphism and a Hamiltonian constraint, resp. two Hamiltonian constraints, leads to a diffeomorphic spacetime.

We also observe that the canonical variables appear on the right hand side in (1.29c). This means that we do not have a true Lie algebra, for which only structure *constants* would be allowed. It will thus be difficult to find representations, and attempts at quantizing this algebra (replacing Poisson brackets by commutators and promoting the canonical variables to operators) will suffer from ordering ambiguities. Note however that the appearance of h^{ab} was to be expected. On the left hand side we have the commutator of two Hamiltonian constraints, that generate evolution in time. The result of the action of $X_{H(N)}$, resp. $X_{H(M)}$ will in both cases be an object on Σ_t. According to the interpretation of the constraints, the commutator can thus only correspond to a diffeomorphism constraint. We therefore need a vector antisymmetric in N and M as argument for the constraint. Without using any background structure, we can only construct an antisymmetric one-form, such as $NdM-MdN$. To convert it into a vector, we clearly need the inverse metric!

CHAPTER 2

Gauge theory formulation of general relativity

We will start this chapter by describing our quantization method. The reader should not be discouraged by the abstract language — we will outline how to recover the basic concepts developed in a standard quantum mechanics course, which should help understand the procedure. Also the ideas should become clearer when we apply them in the next chapter. We continue with a brief overview of why all the attempts to quantize general relativity in the ADM-framework failed so far. This provides the motivation for introducing new variables. The gauge theory formulation will be motivated from the ADM-framework in section 3, the proof that both are equivalent will be given in the following section.

1. The quantization program

Quantum field theories on a fixed background are built on the Wightman axioms. The most important one of these is the requirement of micro-causality: if two points are spacelike separated, one demands that the corresponding operators commute,

$$[\phi(x), \phi(y)] = 0, \quad \forall\, \eta_{ab}(x-y)^a(x-y)^b > 0, \tag{2.1}$$

where η_{ab} denotes the fixed background metric. But in general relativity the metric itself represents the field, we do not have a preferred background! Thus not even the premise of the previous equation does make any sense — we have to use a different procedure for quantization, which we now describe.

As mentioned in the introduction, our preferred quantization method is based on a classical Hamiltonian formulation. We will therefore assume that we are given a phase space (symplectic manifold) $(M, \{\cdot,\cdot\})$, together with a set of first class constraints ϕ_i (including the Hamiltonian). We will also assume that the symplectic manifold is a cotangent bundle $T^*\mathcal{C}$ of a configuration space $\mathcal{C}$, since we are only interested in this case. The Poisson brackets are given in the standard form $\{q^i, p_j\} = \delta^i_j$. Quantization then follows the following procedure:

(1) A famous theorem due to Groenewald and van Hove states that it is impossible to find a quantization of the full Poisson algebra $(C^\infty(M), \{\cdot,\cdot\})$ of all functions on M. We will therefore start by choosing a Poisson subalgebra $\mathcal{A}$, whose elements are to be quantized. The idea to find a quantized version of an observable that is not in $\mathcal{A}$ is then to express it as a function of the basic variables in $\mathcal{A}$ (or at least as a limit of such functions). An important requirement is therefore that all functions on $T^*\mathcal{C}$ can be expressed as limits of functions in $\mathcal{A}$. By the Stone-Weierstrass theorem D.3

this is always the case (at least for all continuous functions) if $\mathcal{A}$ separates the points of $T^*\mathcal{C}$.

For our case of a cotangent bundle $T^*\mathcal{C}$, a natural choice for $\mathcal{A}$ is at hand. Vector fields on the *base* manifold $\mathcal{C}$ are naturally embedded into the functions on $T^*\mathcal{C}$ via $v^a \mapsto p_i v^i$. Here $v^i \colon \mathcal{C} \to \mathbb{R}$ denote the component functions of the vector field v^a in the coordinates q^i. A pair (f, v^a), where f is a smooth function on $\mathcal{C}$, and v^a a vector field on $\mathcal{C}$ thus defines a function on phase space $T^*\mathcal{C}$,

$$F_{(f,v^a)}(q^i, p_i) := f(q^i) + p_j v^j(q^i).$$

The fact that these functions form a vector subspace is obvious. That they also form a Poisson subalgebra can be seen from

$$\begin{aligned}\left\{F_{(f,v^a)}, F_{(g,w^a)}\right\} &= \{f + p_i v^i, g + p_j w^j\} \\ &= \{f, g\} + \{f, p_j w^j\} + \{p_i v^i, g\} + \{p_i v^i, p_j w^j\} \\ &= (w^j \partial_j f - v^j \partial_j g) + (w^j \partial_j v^i - v^j \partial_j w^i) p_i \\ &= F_{(w(f) - v(g), [v,w]^a)},\end{aligned}$$

by formula (C.7). In the last line $v(g)$ denotes the action of a vector field on a function.

For a cotangent bundle, the natural choice is therefore

$$\mathcal{A} := \left\{F_{(f,v^a)} \mid f \in C^\infty(\mathcal{C}), v^a \in \Gamma(T\mathcal{C})\right\}, \tag{2.2}$$

but the quantization procedure also allows more general Poisson subalgebras. In particular we can take the Poisson subalgebra (2.2) but restrict the functions f to a subalgebra of $C^\infty(\mathcal{C})$, as long as this subalgebra is left invariant by the action $\{\cdot, p_i\}$ (this guarantees that $\mathcal{A}$ is a Poisson subalgebra). In LQG for example we allow only so called cylindrical functions.

(2) Now we have to represent the elements of our chosen Poisson subalgebra $\mathcal{A}$ as linear operators on a kinematical Hilbert space $\mathcal{H}_{\text{kin}}$. In other words, we are looking for a linear map $\pi \colon \mathcal{A} \to \mathcal{L}(\mathcal{H}_{\text{kin}})$ and demand that Poisson brackets are represented by commutators,[1]

$$i\hbar\pi\left(\{F, G\}\right) = \left[\pi(F), \pi(G)\right]. \tag{2.3}$$

The typical case is familiar from ordinary quantum mechanics. If $\mathcal{H}_{\text{kin}} = L_2(\bar{\mathcal{C}}, d\mu)$ with (a distributional extension of) the configuration space $\mathcal{C}$ and a probability measure μ thereon, the elements of $\mathcal{H}_{\text{kin}}$ are called wave functions and usually written as $\psi(q)$. The representation map π is often denoted by a hat $\widehat{\ }$ and the

[1]To avoid possible confusion, we note that the representation is a homeomorphism between the Poisson algebra $\mathcal{A}$ and $\mathcal{L}(\mathcal{H}_{\text{kin}})$, which is another Poisson algebra. That is, neither of them is an associative algebra. This construction can be formalized by using the concept of the universal enveloping algebra, as outlined in [40]. Algebraically interested readers are encouraged to take a look at the corresponding pages of wikipedia for an explanation of the concept of universal enveloping algebra.

operator corresponding to configuration and momentum act as usual,

$$\widehat{q}\psi(q) = q\psi(q), \quad \widehat{p}\psi(q) = \frac{\hbar}{i}\frac{\partial}{\partial q}\psi(q).$$

For our preferred Poisson subalgebra for cotangent bundles, this amounts to

$$\widehat{F}_{(f,v^a)}\psi(q^i) = f(q^i)\psi(q^i) + i\hbar v(\psi)(q^i) = f(q^i)\psi(q^i) + i\hbar v^i\partial_i\psi(q^i), \tag{2.4}$$

so functions on $\mathcal{C}$ act by multiplication, whereas vector fields (which correspond to phase space functions linear in the momentum coordinates) act as derivations. The representation property

$$i\hbar\widehat{\{F,G\}} = [\widehat{F},\widehat{G}]$$

is easily verified.

(3) It will usually be a hard task to solve the constraints at the classical level. We therefore define operators $\widehat{\phi}_I$ corresponding to the constraints[2] and try to solve them after quantizing.

The constraints will usually not be defined on all of $\mathcal{H}_{\text{kin}}$, but we assume that there is a dense subset $\mathcal{D}_{\text{kin}}$ on which of all the constraints are defined. In general there will only be distributional solutions of the equation $\widehat{\phi}_I\psi = 0$. To make this precise, we will look for solutions l in the algebraic dual $\mathcal{D}^*_{\text{kin}}$ of $\mathcal{D}_{\text{kin}}$, i.e. the space of distributions. (We could also take the topological dual, but we choose the more general option). We thus have a so called rigged Hilbert space

$$\mathcal{D}_{\text{kin}} \hookrightarrow \mathcal{H}_{\text{kin}} \hookrightarrow \mathcal{D}^*_{\text{kin}},$$

where $\hookrightarrow$ means that the subspace is continuously embedded.

By definition, an operator acts on a distribution by letting the adjoint operator act on the test function. Our constraint equations therefore are

$$(\widehat{\phi}'_I l)(\psi) := l(\widehat{\phi}^\dagger_I\psi) = 0, \quad \forall\,\psi \in \mathcal{D}_{\text{kin}}. \tag{2.5}$$

We denote the space of solutions to all the constraint equations by $\mathcal{D}^*_{\text{phys}}$.

Let us also mention that we can run into problems when the constraint algebra $\{\phi_I,\phi_J\} = f_{IJ}{}^K\phi_K$ contains structure *functions* $f_{IJ}{}^K$, that depend on the phase space coordinates, instead of constants. In that case, they have to be quantized as well and we have

$$[\widehat{\phi}_I,\widehat{\phi}_J] = i\hbar\widehat{\phi}_K\widehat{f}_{IJ}{}^K = i\hbar\left([\widehat{\phi}_K,\widehat{f}_{IJ}{}^K] + \widehat{f}_{IJ}{}^K\widehat{\phi}_K\right). \tag{2.6}$$

Consequently, *every* $l \in \mathcal{D}^*_{\text{phys}}$ also satisfies $\left([\widehat{\phi}_K,\widehat{f}_{IJ}{}^K]\right)' l = 0$. In other words, as long as this commutator is not a linear combination of constraints, the dimension

[2]Remember that in order to define constraint operators we have to express the constraints in terms of the basic variables in $\mathcal{A}$. Unfortunately this might be possible in different ways, so we have to deal with quantization ambiguities.

of $\mathcal{D}_{\text{phys}}$ is reduced and the quantum theory has less physical degrees of freedom, see [38].

(4) Since $\mathcal{D}^*_{\text{phys}}$ is not contained in the kinematical Hilbert space, it does not carry an inner product. The way to regain a physical Hilbert space is to start by defining observables. An operator on $\mathcal{H}_{\text{kin}}$ is called a *strong* Dirac observable, if $[\widehat{O}, \phi_I] = 0\ \forall I$, and a *weak* Dirac observable if $\widehat{O}'\mathcal{D}^*_{\text{phys}} \subseteq \mathcal{D}^*_{\text{phys}}$.

A physical Hilbert space $\mathcal{H}_{\text{phys}}$ is then obtained by taking a subset of $\mathcal{D}^*_{\text{phys}}$ such that all observables are defined on a dense domain $\mathcal{D}_{\text{phys}}$, and equipping this subset with an inner product, such that all observables become self-adjoint operators. We therefore get another rigged Hilbert space

$$\mathcal{D}_{\text{phys}} \hookrightarrow \mathcal{H}_{\text{phys}} \hookrightarrow \mathcal{D}^*_{\text{phys}}.$$

The reader who still feels uneasy about the quantization procedure is strongly encouraged to take a look at Thiemann's lecture notes [38] and quantize the simple systems given as exercises at the end of part I.

The quantum theory of gravity in the ADM formulation was first considered in the 1960-s by DeWitt in his famous series of papers [21, 22, 23]. Unfortunately, the Poisson brackets

$$\{h_{ab}(x), \widetilde{\pi}^{cd}(y)\} = \delta_a^{(c}\delta_b^{d)}\widetilde{\delta}(x-y)$$

are distributional, hence the corresponding operators will be distributions. Now, writing out the Hamiltonian constraint

$$H(N) = \int N\left[\left(\widetilde{\pi}^{ab}\widetilde{\pi}^{cd}h_{ac}h_{bd} - \frac{1}{2}\widetilde{\pi}^{ab}\widetilde{\pi}^{cd}h_{ab}h_{cd}\right)\omega_h^{-1} - {}^3R\omega_h\right],$$

we see that it involves multiplication of distributions, which is not defined. Furthermore, ${}^3R$ and ω_h^{-1} are highly non-linear functions of the canonical variables. One therefore obtained at most formal solutions and was stuck with the quantization of the ADM formulation for a long time.

New hope arose in the 1980-s when Abhay Ashtekar realized that general relativity can be written as a gauge theory, where all constraints take a polynomial form. The new theory is presented in pedagogical form in [4]. Although not up to date anymore, his book is still very much worth reading.

2. Foundations of gauge theories

Gauge theories are formulated in terms of a connection on a principal fiber bundle. Unfortunately, a proper introduction into fiber bundles and connections thereon consumes quite some time and space. Still, most concepts can be introduced without using the notion of fiber bundle by just working on the base manifold of the bundle. Since we want the reader not familiar with fiber bundle theory to be able to follow our treatment, we choose this approach and present the required formulas and definitions. They might look artificial at first sight, but we remind the reader that all the concepts come naturally

when introduced on the bundle. The complicated-looking formulas of this section only arise when we introduce a special trivialization, in physicists terms a special gauge. For a proper discussion of the subject, we refer to the standard literature, [20] and [27].[3]

We will thus define a connection (or gauge potential) A^i_a as a one-form on the manifold M with values in the Lie algebra $\mathfrak{g}$ of some Lie group G. Since we will only be concerned with $G = SO(3)$ or its double cover $SU(2)$, we will restrict ourselves to this case.

We denote the basis elements of the Lie algebra by τ_i, with structure constants $[\tau_i, \tau_j] = \varepsilon_{ij}{}^k \tau_k$. For $\mathfrak{so}(3)$ we use the basis $(\tau_k)^i{}_j = \varepsilon^i{}_{kj}$. From now on, we can therefore freely switch between components v^k with respect to the basis, and matrix elements $v^i{}_j = v^k \varepsilon^i{}_{kj}$ in the fundamental representation.

Given a connection, we can extend the covariant derivative D_a to tensors with values in the Lie algebra $\mathfrak{g}$. We define the action of D_a on $\mathfrak{g}$-valued scalars as

$$D_a v^i := \partial_a v^i + A_a{}^i{}_j v^j.$$

The action on arbitrary $\mathfrak{g}$-valued tensors is then already defined by the axioms for a derivative operator (linearity, Leibniz rule and compatibility with contractions) and the known action on ordinary tensors.[4]

A connection also defines an extension of the exterior derivative d to Lie algebra valued *forms* ϕ by

$$D\phi := d\phi + [A, \phi]\,. \tag{2.7}$$

Unfortunately this derivative is also denoted by the same letter D. We agree that we shall mean the extension of the exterior derivative, when it acts on forms only and also whenever we write D without abstract indices (dropping the indices of forms). On the other hand, if D_a acts on a vector index we shall mean the extension of the covariant derivative (exceptions will be stated explicitly). Note that both derivatives agree on Lie

[3]The situation probably can be compared to the way tensors are introduced in most beginning physicists courses. The definition of a tensor as an object that "transforms like a tensor" suffices to do a lot of calculations. On the other hand, real insight is only possible once one is given the proper definition as a multilinear map.

[4]To find the action on ω^j_b, for example, one can calculate for arbitrary vectors v^b

$$(D_a - \partial_a)\left(\omega^j_b v^b\right) = A_a{}^j{}_k \omega^k_b v^b.$$

On the other hand,

$$(D_a - \partial_a)\left(\omega^j_b v^b\right) = (D_a - \partial_a)\,\omega^j_b v^b + \omega^j_b \Gamma^b{}_{ac} v^c.$$

By comparison we obtain

$$D_a \omega^j_b = \partial_a \omega^j_b - \Gamma^c{}_{ab}\omega^j_c + A_a{}^j{}_k \omega^k_b.$$

The generalization to arbitrary Lie algebra valued tensors is straightforward:

$$D_a t^{b\ldots}{}_{c\ldots}{}^i = \mathring{D}_a t^{b\ldots}{}_{c\ldots}{}^i + A_a{}^i{}_j t^{b\ldots}{}_{c\ldots}{}^j,$$

where $\mathring{D}$ acts only on tensorial indices, in other words it treats the index i as a counting index. The additional term can also be written in the form (outlined here only for a scalar)

$$[A, v]^i = A^k v^j\,[\tau_k, \tau_j]^i = \varepsilon^i{}_{kj} A^k v^j = A^i{}_j v^j.$$

Finally, consider $D_a(v^i w_i)$ to find that for a lower Lie algebra index the additional term is given by $-A^j{}_i w_j$.

algebra valued scalars, and also that the derivative of forms can be obtained by taking the derivative of the previous paragraph and antisymmetrizing in all indices.[5]

We will now briefly recapitulate the vielbein calculus, since it provided the primary inspiration for introducing a gauge theory formulation of general relativity and also because we can use it to motivate the concepts of this chapter. We define cotetrads (in the present three-dimensional context they are called cotriads) by

$$h_{ab} = \delta_{ij} e^i_a e^j_b. \tag{2.8}$$

The triad is then given by $E^a_i e^j_a = \delta^j_i$. An extension of the exterior derivative to vector-valued differential forms is can be defined via

$$(d^\nabla \omega^a)_{I_{p+1}} := \frac{1}{p!} \delta^{bI_p}_{I_{p+1}} \nabla_b \omega^a_{I_p}.$$

Contrary to its action on ordinary differential forms this definition *does* depend on the choice of ∇ (for the purpose of this paragraph ∇ denotes a derivation on Σ). Acting with this exterior derivative on a basis vector E^a_i we obtain a vector-valued one-form which can be decomposed in the basis E^a_i,

$$d^\nabla E^a_i =: \omega^j{}_i E^a_j. \tag{2.9}$$

The components $\omega^i{}_j$ are one-forms and called Ricci-rotation coefficients.

Interpreting $\delta^a_b = \partial^a_i dx^i_b = E^a_i e^i_b$ as vector-valued one-form, we can deduce the first structure equation. On the one hand we have

$$(d^\nabla \delta^a)_{bc} = 2\nabla_{[c}\partial^a_i dx^i_{b]} + 2\partial^a_i (d^2 x^i)_{bc} = 2\partial_{[c}\partial^a_i dx^i_{b]} + 2\Gamma^a{}_{d[c}\partial^d_i dx^i_{b]} = 2\Gamma^a{}_{ij} dx^j_{[c} dx^i_{b]} = 0.$$

On the other hand we also have

$$d^\nabla \delta^a = d^\nabla E^a_i e^i + E^a_i de^i = \omega^j{}_i E^a_j \wedge e^i + E^a_i de^i = \left(\omega^j{}_i \wedge e^i + de^j\right) E^a_j.$$

Comparison therefore yields

$$de^i + \omega^i{}_j \wedge e^j = 0. \tag{2.10}$$

By acting on $h^{ab} = \delta^{ij} E^a_i E^b_j$ with the exterior derivative d^∇ (where now we have to use the derivative compatible with h_{ab} for ∇), we easily see that ω^{ij} is antisymmetric, that is we can interpret the Ricci-rotation coefficients as one-form with values in the Lie-algebra $\mathfrak{so}(3)$. According to the above definition, $\omega^i{}_j$ thus defines a connection, which we will call Levi Civita connection (sometimes it is also referred to as spin connection). We will now examine why this connection plays a special role.

Since E^a_i is an isomorphism between $\mathfrak{g}$ and the tangent space at a given point, it is natural to consider those connections that satisfy

$$D_a E^b_i = 0, \tag{2.11}$$

[5]Since the exterior derivative on ordinary forms is given even in the absence of a notion of parallel transport, we can define this type of derivative in any case, whereas the derivative from the previous paragraph requires additional structure to be present!

because then we have for example $D_a\left(v^i E^b_i\right) = D_a v^i E^b_i$, i.e. differentiation is compatible with the isomorphism provided by E^a_i. Writing out equation (2.11) yields (we denote by $\mathring{D}$ the derivation that only acts on tensor indices)

$$\mathring{D}_a E^b_i - A_a{}^j{}_i E^b_j = \left(d^\nabla E^b_i - A^j{}_i E^b_j\right)_a = 0,$$

which is precisely solved by the Levi-Civita connection.[6]

For the Levi Civita connection we know that the Riemann tensor can be calculated via the second structure equation

$$R^i{}_j = d\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j. \tag{2.12}$$

It is then easy too see that $D^2\phi = [R, \phi]$ for arbitrary $\mathfrak{g}$-valued scalars ϕ. We take this as our motivation to define the curvature two-form F^i_{ab} of an arbitrary connection A^i_a via

$$D^2\phi =: [F, \phi]\,. \tag{2.13}$$

A short calculation using (N.19) and (N.18),

$$\begin{aligned} D^2\phi = D\left(d\phi + [A,\phi]\right) &= d^2\phi + [dA,\phi] - [A, d\phi] \\ &+ [A, d\phi] + [A,[A,\phi]] = \left[dA + \frac{1}{2}[A,A]\,, \phi\right], \end{aligned}$$

then yields the following expression for the curvature,

$$F = dA + \frac{1}{2}[A,A]\,. \tag{2.14}$$

Finally we introduce the concept of a gauge transformation, which is induced by a function $g : M \to G$. The connection transforms under this change of gauge via

$$A \to {}^gA = g^{-1}dg + g^{-1}Ag. \tag{2.15}$$

We will often work with infinitesimal gauge transformations,

$$g = \mathbb{1} + \Lambda$$

where Λ takes values in $\mathfrak{g}$. The transformation of the connection then reads (we neglect higher powers of Λ in the first-order change)

$${}^gA = (\mathbb{1} - \Lambda)d(\mathbb{1} + \Lambda) + (\mathbb{1} - \Lambda)A(\mathbb{1} + \Lambda) = d\Lambda + [A, \Lambda] + A = A + D\Lambda.$$

3. New variables for general relativity

We have just seen that the $\mathfrak{so}(3)$-valued one-form ω^i_a defines a connection, and it would be natural to guess that $(\omega^i_a, \widetilde{E}^a_i)$ form a canonically conjugate pair. However, the definition (2.9) (or equivalently the structure equation (2.10)) tells us that ω^i_a is determined entirely by $\widetilde{E}^a_i$, therefore $\{\omega^i_a, \widetilde{E}^b_j\} = 0$. Taking a look at the ADM formulation

[6]For the extension of the exterior derivative we have the analog compatibility requirement $De^i = 0$ for the inverse isomorphism. Of course this requirement also uniquely singles out the Levi Civita connection, since it is equivalent to (2.10) which uniquely determines $\omega^i{}_j$.

it is natural to consider the extrinsic curvature. We observe that we can decompose the first index of $K^a{}_b$ in an orthonormal basis,

$$K^a{}_b = K^i_b E^a_i. \tag{2.16}$$

Since a change of the orthonormal basis corresponds to a rotation, the index i on K^i_b transforms again in the Lie algebra $\mathfrak{so}(3)$. We can therefore define a new connection by a linear combination of ω^i_a and K^i_a. For our phase space we choose the symplectic manifold with Darboux coordinates

$$A^i_a := \omega^i_a + \beta K^i_a, \quad \widetilde{E}^a_i := E^a_i \omega_h, \tag{2.17}$$

where $\beta \neq 0$ is an unspecified parameter called the "Immirzi parameter". Since we now have two different connections in use, we agree that we denote the gauge-covariant derivative corresponding to ω^i_a by D and the derivative corresponding to the Ashtekar connection A^i_a by $\mathcal{D}$.

In order to show that this phase space properly describes gravitation, we have to show that h_{ab} and $\widetilde{\pi}^{ab}$, when expressed as functions of A^i_a and $\widetilde{E}^a_i$, form a canonical pair as long as A^i_a and $\widetilde{E}^a_i$ do. We will prove this equivalence in the next section and first deal with a more pressing issue.

3.1. Gauß constraint. A^i_a and $\widetilde{E}^a_i$ together have 18 degrees of freedom, whereas the ADM variables have only 12. This means that in the new formulation we have to introduce an additional constraint. Since all triads give the same metric when rotated, this was to be expected and we also know what to do — we have to implement invariance under rotations, in other words invariance under $\mathfrak{so}(3)$ gauge transformations.

We therefore demand that for any function f on phase space we have $f(A + \mathcal{D}\Lambda) = f(A)$, where Λ is an infinitesimal gauge transformation. To see what this implies, we use the following generalization of the first Taylor polynomial to infinite dimensions (as usual we replace sums by integrals and derivatives by the functional derivative)

$$f(A + \delta A) = f(A) + \int \delta A^i_a \frac{\delta f}{\delta A^i_a} = f(A) + \int \left\{ f, \delta A^i_a \widetilde{E}^a_i \right\} = f(A) + \left\{ f, \int \delta A^i_a \widetilde{E}^a_i \right\}.$$

We therefore have (remember that exterior and covariant derivative agree on Λ)

$$f(A + \mathcal{D}\Lambda) = f(A) + \left\{ f, \int \widetilde{E}^a_i (\mathcal{D}\Lambda)^i_a \right\} = f(A) - \left\{ f, \int \Lambda^i \mathcal{D}_a \widetilde{E}^a_i \right\},$$

by a partial integration and assuming compact support of the function Λ to be able to drop boundary terms.

Invariance under rotations therefore enforces the so called "Gauß constraint"

$$\int \Lambda^i \mathcal{D}_a \widetilde{E}^a_i = 0, \quad \forall \Lambda^i \in C^\infty_0(M) \tag{2.18}$$

3.2. Diffeomorphism constraint. We still have to write the two constraints of the ADM formulation in the new variables. Under a diffeomorphism ϕ of the base manifold,

the connection transforms as $A \to \phi^* A$. Considering a one-parameter family $\phi_\varepsilon = \exp(\varepsilon\xi)$ we obtain the infinitesimal version $\phi_\varepsilon^* A = A + \varepsilon \frac{d}{d\varepsilon}\phi_\varepsilon^* A\big|_{\varepsilon=0} + O(\varepsilon^2)$. Under infinitesimal diffeomorphisms the connection therefore transforms as $A_a^i \to A_a^i + \mathcal{L}_\xi A_a^i$.

We can rewrite this expression using (N.12) and the fact the $\xi \lrcorner$ is an antiderivation,

$$\begin{aligned}\mathcal{L}_\xi A &= \xi \lrcorner dA + d(\xi \lrcorner A) = \xi \lrcorner F - \xi \lrcorner \frac{1}{2}[A, A] + d(\xi \lrcorner A) \\ &= \xi \lrcorner F - \frac{1}{2}[\xi \lrcorner A, A] + \frac{1}{2}[A, \xi \lrcorner A] + d(\xi \lrcorner A) = \xi \lrcorner F + \mathcal{D}(\xi \lrcorner A).\end{aligned}$$

Using again our formula for the Taylor polynomial we get

$$f(A + \mathcal{L}_\xi A) = f(A) + \left\{ f, \int \widetilde{E}_i^a \xi^b F_{ba}^i \right\} - \left\{ f, \int \xi^b A_b^i \mathcal{D}_a \widetilde{E}_i^a \right\},$$

where the last term could be treated analog to the Gauß constraint because $\xi \lrcorner A$ is again a Lie algebra valued function and therefore defines a gauge transformation.[7]

Assuming that the Gauß constrained is already solved, this term vanishes and we can write the diffeomorphism constraint as

$$V(\xi) = \int \xi^a \widetilde{E}_i^b F_{ba}^i = 0, \quad \forall \xi \in \Gamma(TM). \tag{2.19}$$

3.3. Hamiltonian constraint. We now consider the Hamiltonian constraint, which is at the core of the Ashtekar variables. The corresponding constraint in the ADM variables (1.11) contains the curvature scalar ${}^3R = E_j \lrcorner (E_i \lrcorner R^{ij})$. Let us now consider the "curvature scalar" of F^{ij}, i.e.

$$E_j \lrcorner (E_i \lrcorner F^{ij}), \tag{2.20}$$

where $F^{ij} := F^k \varepsilon^i{}_k{}^j$. As already explained before, this just means that we go over from calculating with components of $\mathfrak{so}(3)$-valued objects in a certain basis to using matrix elements in the fundamental representation.

Inserting the definition of the connection (2.17), we see that (use (N.17) to obtain the Levi Civita derivative in the second term)

$$F^{ij} = R^{ij} + \beta(DK)^{ij} + \frac{\beta^2}{2}[K, K]^{ij}. \tag{2.21}$$

We will calculate the inner product with E_i and E_j of the three terms separately. The first term obviously gives ${}^3R$, for the second we use the compatibility condition (2.11) for the Levi Civita connection to project the one-form indices directly onto the basis vectors E_i,

$$\begin{aligned}E_j \lrcorner (E_i \lrcorner DK^{ij}) &= E_j \lrcorner D_i K^{ij} = D_i K_j{}^{ij} \\ &= D_i K_j{}^k \varepsilon^i{}_k{}^j = D_i K^{jk} \varepsilon^i{}_{kj} = 0\end{aligned}$$

provided that K_{ab} is symmetric, in other words provided that the Gauß constraint holds (see remark 2.1).

[7]Observe that at this point there a partial integration involved. The corresponding boundary terms can be dropped only if $\xi \lrcorner A$ has compact support. The diffeomorphism constraint therefore only enforces invariance under *local* diffeomorphism.

The third term can be written as

$$[K,K]^{ij} = (K^m \wedge K^n)\,[\tau_m,\tau_n]^{ij} = (K^m \wedge K^n)\,\varepsilon_{mn}{}^{l}\tau_l^{ij} = (K^m \wedge K^n)\,\varepsilon_{mn}{}^{l}\varepsilon^{i\,j}_{\;\;l}.$$

We obtain by the anti-derivation property of $\xi\lrcorner$

$$E_i \lrcorner [K,K]^{ij} = (K_i^m K^n - K^m K_i^n)\,\varepsilon_{mn}{}^{l}\varepsilon^{i\,j}_{\;\;l},$$

and further

$$\begin{aligned} E_j \lrcorner \left(E_i \lrcorner [K,K]^{ij}\right) &= \left(K_i^m K_j^n - K_j^m K_i^n\right)\varepsilon_{mn}{}^{l}\varepsilon^{i\,j}_{\;\;l} = 2K_i^m K_j^n \varepsilon_{mn}{}^{l}\varepsilon^{i\,j}_{\;\;l} \\ &= 2\left(K_i^j K_j^i - K_i^i K_j^j\right) = -2\left(K^2 - K_{ab}K^{ab}\right). \end{aligned}$$

Collecting terms, we have

$$E_j \lrcorner \left(E_i \lrcorner F^{ij}\right) = {}^3R - \beta^2\left(K^2 - K_{ab}K^{ab}\right).$$

For the special cases $\beta = \pm i$, the right hand side equals (minus) the Hamiltonian constraint of the ADM-formulation and we get the simple result

$$\widetilde{\widetilde{H}}_E = \widetilde{E}_i^a \widetilde{E}_j^b F_{ab}^k \varepsilon^{ij}{}_k.$$

The index E on the Hamiltonian constraint for the value $\beta = \pm i$ stems from the fact that this is also the correct expression for the case of a theory with positive definite four-dimensional metric.

For arbitrary values of the Immirzi parameter we get a correction term

$$\begin{aligned} H - H_E &= -(1+\beta^2)\left(K^2 - K_{ab}K^{ab}\right) \\ &= -(1+\beta^2)\left(K_a^i K_b^j - K_a^j K_b^i\right) E_i^a E_j^b \\ &= -\left(1+\beta^2\right) K_a^k K_b^l \delta_{kl}^{ij} E_i^a E_j^b \\ &= -(1+\beta^2)\varepsilon_{mkl} K_a^k K_b^l \widetilde{E}_i^a \widetilde{E}_j^b \epsilon^{mij}\omega_h^{-2}, \end{aligned}$$

and the full Hamiltonian constraint reads

$$\widetilde{H} = \varepsilon^{mij}\widetilde{E}_i^a \widetilde{E}_j^b \left(F_{ab}^m - (1+\beta^2)\varepsilon_{mkl}K_a^k K_b^l\right)\omega_h^{-1}. \tag{2.22}$$

This expression looks like we have not gained much, but we will see in section 5.2 that this is not the case.

4. Equivalence of gauge theory- and ADM-formulation

Summarizing our findings, general relativity can be described by a phase space with Darboux coordinates $(A_a^i, \widetilde{E}_j^b)$, satisfying (the factor $\frac{\beta}{2}$ is conventional and some authors choose to get rid of it by redefining $\widetilde{E}_j^a$)

$$\left\{A_a^i(x), A_b^j(y)\right\} = \left\{\widetilde{E}_i^a(x), \widetilde{E}_j^b(y)\right\} = 0, \quad \left\{A_a^i(x), \widetilde{E}_j^b(y)\right\} = \frac{\beta}{2}\widetilde{\delta}(x-y)\delta_a^b\delta_j^i. \tag{2.23}$$

The constraints of the theory are

$$\text{“Gauß”} \qquad \widetilde{G}_i = \mathcal{D}_a \widetilde{E}^a_i \tag{2.24}$$

$$\text{“Diffeomorphism”} \qquad \widetilde{V}_a = \widetilde{E}^b_i F^i_{ba} \tag{2.25}$$

$$\text{“Hamiltonian”} \qquad \widetilde{H} = \varepsilon^{mij} \widetilde{E}^a_i \widetilde{E}^b_j \left(F^m_{ab} - (1+\beta^2)\varepsilon_{mkl} K^k_a K^l_b \right) \omega_h^{-1}, \tag{2.26}$$

where $\mathcal{D} := d + [A, \cdot]$ is the gauge covariant derivative and $F = dA + \frac{1}{2}[A, A]$ the curvature of the connection A.

In order to show that this formulation is equivalent to the ADM framework, we have to express the ADM variables h_{ab} and $\widetilde{\pi}^{ab}$ as functions of the new variables via the defining equations (2.17) and then show that the so-defined functions also satisfy canonical Poisson brackets, provided that the Gauß constraint holds.

4.1. $\{h_{ab}, h_{cd}\}$**.** We start by observing that the volume element can be expressed solely in terms of the densitized triad via

$$\omega_h = \sqrt{\frac{1}{3!} \underset{\sim}{\delta}_{abc} \varepsilon^{ijk} \widetilde{E}^a_i \widetilde{E}^b_j \widetilde{E}^c_k}. \tag{2.27}$$

For the definition of $\underset{\sim}{\delta}_{abc}$ see the notations section. To verify this identity note that

$$\frac{1}{3!} \underset{\sim}{\delta}_{abc} \varepsilon^{ijk} \widetilde{E}^a_i \widetilde{E}^b_j \widetilde{E}^c_k = \omega_h^3 \frac{1}{3!} \underset{\sim}{\delta}_{abc} \varepsilon^{abc} = \omega_h^3 \omega_h^{-1} = \omega_h^2.$$

Now ω_h and $\widetilde{E}^a_i$ together determine E^a_i via (2.17). The cotriad e^i_a is dual to E^a_i and determines h_{ab} via (2.8). We conclude that h_{ab} is a function of $\widetilde{E}^a_i$ only and does not depend on A^i_a. By the formula (C.8) for Poisson brackets in Darboux coordinates we get

$$\{h_{ab}, h_{cd}\} = 0.$$

4.2. $\left\{h_{ab}, \widetilde{\pi}^{cd}\right\}$**.** Since we just saw that h_{ab} does not depend on A^i_a we just need the variations $\delta_{\widetilde{E}} h_{ab}$ and $\delta_A \widetilde{\pi}^{ab}$.

Let us start with the first one. By the definition (2.8) of the triad,

$$\delta_{\widetilde{E}} h_{ab} = \delta e^i_a \delta_{ij} e^j_b + e^i_a \delta_{ij} \delta e^j_b. \tag{2.28}$$

We get the variation of the cotriad from the duality relation $e^i_a E^a_j = \delta^i_j$,

$$\delta e^i_a E^a_j + e^i_a \delta E^a_j = 0,$$

which gives after multiplication with e^j_b

$$\delta e^i_b = -e^i_a \delta E^a_j e^j_b. \tag{2.29}$$

Renaming indices and using $\delta E^c_k = \delta \widetilde{E}^c_k \omega_h^{-1} + \widetilde{E}^c_k \delta(\omega_h^{-1}) = \delta \widetilde{E}^c_k \omega_h^{-1} - \widetilde{E}^c_k \omega_h^{-2} \delta \omega_h$ leads to

$$\delta e^i_a = -e^i_c \delta \widetilde{E}^c_k \omega_h^{-1} e^k_a + e^i_c E^c_k \omega_h^{-1} \delta \omega_h e^k_a. \tag{2.30}$$

To express the variation of the volume form occurring in the last term we use equation (B.2),

$$\begin{aligned}\delta\omega_h &= \frac{1}{2}h^{ab}\delta h_{ab}\omega_h = \frac{1}{2}E^a_i E^b_i \delta\left(e^j_a e^j_b\right)\omega_h = \frac{1}{2}2E^a_i\delta e^i_a\omega_h \\ &= -E^a_i e^i_b \delta E^b_j e^j_a \omega_h = -\delta E^a_i e^i_a \omega_h \\ &= -\delta\widetilde{E}^a_i e^i_a + E^a_i e^i_a \delta\omega_h = -\delta\widetilde{E}^a_i e^i_a + 3\delta\omega_h,\end{aligned}$$

by using (2.29) to go from the first to the second line and applying the Leibniz rule from the second to the third. Subtracting the last term we obtain

$$\delta\omega_h = \frac{1}{2}e^i_a\delta\widetilde{E}^a_i. \tag{2.31}$$

Inserting into (2.30) we get

$$\delta e^i_a = -e^i_c\delta\widetilde{E}^c_k e^k_a\omega_h^{-1} + \frac{1}{2}e^i_a e^k_c\delta\widetilde{E}^c_k\omega_h^{-1}.$$

Using this expression in (2.28) yields

$$\begin{aligned}\delta_{\widetilde{E}}h_{ab} &= -h_{bc}\delta\widetilde{E}^c_k e^k_a\omega_h^{-1} + \frac{1}{2}h_{ab}\delta\widetilde{E}^c_k e^k_c\omega_h^{-1} - h_{ac}\delta\widetilde{E}^c_k e^k_b\omega_h^{-1} + \frac{1}{2}h_{ab}\delta\widetilde{E}^c_k e^k_c\omega_h^{-1} \\ &= -h_{bc}\delta\widetilde{E}^c_k e^k_a\omega_h^{-1} - h_{ac}\delta\widetilde{E}^c_k e^k_b\omega_h^{-1} + h_{ab}\delta\widetilde{E}^c_k e^k_c\omega_h^{-1}\end{aligned} \tag{2.32}$$

For the variation $\delta_A\widetilde{\pi}^{ab}$ we use expression (1.15),

$$\begin{aligned}\widetilde{\pi}^{ab} &= \left(K^{ab} - h^{ab}K\right)\omega_h = \left(K^{ib}E^a_i - h^{ab}K^i_c E^c_i\right)\omega_h \\ &= K^i_c\left(E^a_i h^{bc}\omega_h - h^{ab}E^c_i\omega_h\right) = K^i_c\left(\widetilde{E}^a_i h^{bc} - h^{ab}\widetilde{E}^c_i\right).\end{aligned} \tag{2.33}$$

From the definition (2.17) of the connection we get $\delta_A A^i_a = \beta\delta_A K^i_a$, because ω^i_a does not depend on A^i_a. Since, as already remarked several times, also h_{ab} does not depend on the connection, we get

$$\delta_A\widetilde{\pi}^{cd} = \frac{1}{\beta}\delta A^i_f\left(\widetilde{E}^c_i h^{fd} - \widetilde{E}^f_i h^{cd}\right). \tag{2.34}$$

Due to the extra factor of $\frac{\beta}{2}$ in the Poisson brackets (2.23), our standard formula for Poisson brackets (C.8) acquires an extra factor, hence

$$\begin{aligned}\left\{h_{ab}(x), \widetilde{\pi}^{cd}(y)\right\} &= -\frac{\beta}{2}\int_z\left(-h_{bf}e^k_a\omega_h^{-1} - h_{af}e^k_b\omega_h^{-1} + h_{ab}e^k_f\omega_h^{-1}\right)(z)\widetilde{\delta}(x-z) \\ &\qquad\times\frac{1}{\beta}\left(\widetilde{E}^c_k h^{fd} - \widetilde{E}^f_k h^{cd}\right)(z)\widetilde{\delta}(z-y) \\ &= -\frac{1}{2}\int_z\left(-\delta^d_b\delta^c_a + h_{ab}h^{cd} - \delta^d_a\delta^c_b + h_{ab}h^{cd} + h_{ab}h^{cd} - h_{ab}3h^{cd}\right)(z) \\ &\qquad\times\widetilde{\delta}(x-z)\widetilde{\delta}(z-y) \\ &= \frac{1}{2}\left(\delta^d_a\delta^c_b + \delta^c_a\delta^d_b\right)\int_z\widetilde{\delta}(x-z)\widetilde{\delta}(z-y) \\ &= \delta^{(d}_a\delta^{c)}_b\widetilde{\delta}(x-y),\end{aligned}$$

where we used the convolution of the δ-distributions, $\int_z \widetilde{\delta}(x-z)\widetilde{\delta}(z-y) = \widetilde{\delta}(x-y)$ which can be proved by application to arbitrary test functions φ,

$$\left(\widetilde{\delta}(x-z), \left(\widetilde{\delta}(z-y), \varphi(y)\right)\right) = \left(\widetilde{\delta}(x-z), \varphi(z)\right) = \varphi(x) = \left(\widetilde{\delta}(x-y), \varphi(y)\right).$$

4.3. $\left\{\widetilde{\pi}^{ab}, \widetilde{\pi}^{cd}\right\}$**.** The calculation of the last Poisson bracket is the most laborious one. We will need to make use of the following remark.

Remark 2.1. Since K_{ab} is now defined via equation (2.17), it is a priori not a symmetric tensor. However, it is easy to see that its symmetry, $K_{ab} = K_{ba}$, is equivalent to the Gauß constraint. From the definition (2.17) we have

$$\widetilde{G}_i = \mathcal{D}_a \widetilde{E}^a_j = D_a \widetilde{E}^a_j + \beta \varepsilon_{ij}{}^k K^j_a \widetilde{E}^a_k = \beta \varepsilon_{ij}{}^k K^j_k \omega_h = \beta \varepsilon_i{}^{jk} K_{jk} \omega_h,$$

where the first term of the last expression in the first line vanished because of the compatibility requirement (2.11) on the spin connection. We thus arrive at the statement that

$$\widetilde{G}_i = 0 \Leftrightarrow K_{ab} = K_{ba}.$$

The variation $\delta_A \widetilde{\pi}^{ab}$ was already calculated in (2.34). For the second variation, $\delta_{\widetilde{E}} \widetilde{\pi}^{ab}$ we take expression (2.33) to observe that this task essentially reduces to the calculation of $\delta_{\widetilde{E}} K^i_a$. Remembering the definition (2.17) of the connection, we see that $\delta_{\widetilde{E}} K^i_a = -\frac{1}{\beta} \delta_{\widetilde{E}} \omega^i_a$. This last variation is found by varying the definition of the Levi Civita connection (2.9). We obtain

$$d\delta e^i + \delta\omega^i{}_j \wedge e^j + \omega^i{}_j \wedge \delta e^j = D\delta e^i + \delta\omega^k \varepsilon^i{}_{kj} e^j = 0. \tag{2.35}$$

Writing the Poisson bracket in the form

$$\left\{\widetilde{\pi}^{ab}, \widetilde{\pi}^{cd}\right\} = \frac{\beta}{2} \left(\delta_{\widetilde{E}} \widetilde{\pi}^{ab}\big|_{\delta \widetilde{E}^f_i = \frac{\delta \widetilde{\pi}^{cd}}{\delta A^i_f}} - \delta_{\widetilde{E}} \widetilde{\pi}^{cd}\big|_{\delta \widetilde{E}^f_i = \frac{\delta \widetilde{\pi}^{ab}}{\delta A^i_f}} \right)$$

we see that we only need δK^i_a evaluated at $\delta \widetilde{E}^f_i = \frac{\delta \widetilde{\pi}^{ab}}{\delta A^i_f}$, which according to (2.34) depends only on quantities annihilated by D. Since also δe^i can be expressed by such quantities, we conclude by equation (2.35) that $\delta K^i_a\big|_{\delta \widetilde{E}^f_i = \frac{\delta \widetilde{\pi}^{ab}}{\delta A^i_f}} = 0$.

It therefore remains to calculate $K^i_f \delta_{\widetilde{E}} \left(\widetilde{E}^a_i h^{bf} - \widetilde{E}^f_i h^{ab} \right)\Big|_{\delta \widetilde{E}^f_i = \frac{\delta \widetilde{\pi}^{ab}}{\delta A^i_f}}$, which reduces to carefully inserting already known variations. One may for example write everything in terms of $\widetilde{E}^a_i$ and ω_h and use (2.31) to arrive at

$$\left\{\widetilde{\pi}^{ab}, \widetilde{\pi}^{cd}\right\} = \frac{1}{2} \omega_h \left(K^{ac} - K^{ca}\right) h^{bd},$$

which equals zero provided that the Gauß constraint holds.

This shows that enlargement of the phase space w.r.t. the Ashtekar variables is equivalent to the original ADM formulation as long as we identify locally rotationally equivalent phase space points.

CHAPTER 3

Loop quantum gravity

While in the context of the ADM formulation the Hamiltonian constraint is highly non-linear, we have seen in equation (2.22) that at least for $\beta = \pm i$ the new formulation of GR leads to a polynomial expression. Historically this was the starting point for a new approach to finding a quantum theory of gravity. LQG is the attempt to apply the quantization program of section 1 in the previous chapter to the new formulation of general relativity.

We will see in this chapter that now the quantization program can almost be completed. We start by smearing the connection variables and the densitized triads, thereby arriving at our preferred Poisson subalgebra (3.13). We then proceed by constructing a representation space in section 2. Section 3 explains how to implement the kinematical constraints. We will see that one can construct a rigorously defined space of solutions, which is tremendous progress compared to the ADM formulation. There is also a well-defined quantization of the Hamiltonian constraint available and a large class of solutions can be constructed. Unfortunately it involves quite some ambiguities — we will hint at these in section 5 and point the reader to Thiemann's original article [37] for a more detailed discussion.

We will also discuss operators corresponding to length, area and volume and see that one can calculate (properties of) their spectra. This is a major success of LQG since quantization of space is derived and not postulated.

Our presentation follows Thiemann [40], [38] and Rovelli [32] — we therefore also refer the reader to their treatments for some details and topics that had to be omitted for lack of time and space.

1. The holonomy-flux algebra

Since the Poisson brackets (2.23) are still singular, the canonical variables A^i_a and $\widetilde{E}^a_i$ will correspond to operator-valued distributions in the quantum theory. The most straightforward way to deal with this problem would be to smear them with test-functions.[1] The problem with this choice of smearing is that the complicated behavior under gauge-transformations (2.15) carries over to the smeared connection (see e.g. [40]). It will therefore be difficult to find gauge-invariant observables. Fortunately the problem arises in all

[1]For $\{\phi(x), \widetilde{\pi}(y)\} = \widetilde{\delta}(x-y)$ this amounts to defining $\phi(\widetilde{f}) := \int \phi \widetilde{f}$ and $\widetilde{\pi}(g) := \int \widetilde{\pi} g$. The distributional Poisson brackets are then replaced by $\left\{\phi(\widetilde{f}), \widetilde{\pi}(g)\right\} = \int \widetilde{f} g$.

non-Abelian gauge-theories, allowing us to take advantage of the work that has been done for them.

1.1. Configuration variables. For the smearing of the connection we consider the set $\mathcal{C}$ of piecewise analytic, continuous, oriented, embedded (see section 1 in appendix A) curves $c\colon [0,1] \to \Sigma$. We define the holonomy of a given connection A^i_a along a curve c as the value at $t = 1$ of the unique solution to the differential equation

$$\dot{h}_c(t) + A(t)h_c(t) = 0, \quad h_c(0) = \mathbb{1}, \tag{3.1}$$

for $SU(2)$-valued functions $h_c(t)$, where $A(t) := A^i_a(c(t))\tau_i \dot{c}^a(t)$.[2] Holonomies are denoted by $h_c(1) =: h_c(A) =: A(c)$, depending on whether one wants to emphasize the dependence on the curve or the connection.

Example 3.1. The name holonomy is usually reserved for the parallel transport along closed curves. We use it here also for parallel transport along arbitrary curves c. The definition again comes naturally when introduced in a fiber bundle theoretic context. However, we can give a good motivation by considering the Levi Civita connection. From the basic relativity course we know the defining equation of the parallel transport of a vector v^a along a curve with tangent vector u^a,

$$(u^a D_a)v^b = 0.$$

Writing this equation in coordinates yields (using the chain rule for the first term)

$$\left(\dot{x}^j \partial_j\right) v^k + \Gamma^k{}_{lm}\dot{x}^m v^l = \dot{v}^k + \Gamma^k{}_l v^l = 0$$

with $\Gamma^k{}_l := \Gamma^k{}_{lm}\dot{x}^m$. With an ansatz $v^k(t) = h^k{}_l(t)v^l(0)$ this yields the differential equation for the holonomy $h^k{}_l$,

$$\left(\dot{h} + \Gamma h\right) v(0) = 0,$$

for arbitrary vectors v^a, hence

$$\dot{h} + \Gamma h = 0.$$

In physical terms this looks like a Schrödinger equation for the time evolution operator.

Let us mention some important properties of holonomies that we will need later on. We denote the beginning and end point of a curve by $b(c)$ and $f(c)$, respectively and define the composition of curves with $f(c_1) = b(c_2)$ by

$$(c_2 \circ c_1)(t) := \begin{cases} c_1(2t) & t \in [0, \frac{1}{2}] \\ c_2(2t-1) & t \in [\frac{1}{2}, 1], \end{cases} \tag{3.2}$$

as well as inversion of curves by

$$c^{-1}(t) := c(1-t). \tag{3.3}$$

[2]From now on we work with the double cover $SU(2)$ of $SO(3)$. As generators of the corresponding Lie algebra $\mathfrak{su}(2)$ we take $\tau_i := -\frac{i}{2}\sigma_i$ with the Pauli-matrices σ_i

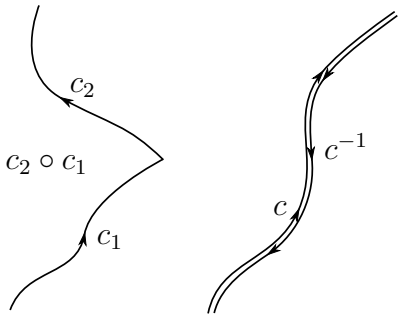

FIGURE 1. Composition of curves with $f(c_1) = b(c_2)$ and inversion of curves.

We then have the important properties

$$\begin{aligned} A(c_2 \circ c_1) &= A(c_2)A(c_1) \text{ and} \\ A(c^{-1}) &= (A(c))^{-1}. \end{aligned} \tag{3.4}$$

The proof is easy and most conveniently done by showing that the functions

$$\begin{aligned} f(t) &:= \begin{cases} h_{c_1}(2t) & 0 \le t < \frac{1}{2} \\ h_{c_2}(2t-1)h_{c_1}(1) & \frac{1}{2} \le t \le 1, \end{cases} \\ g(t) &:= h_c(1-t)(h_c(1))^{-1} \end{aligned}$$

satisfy the differential equations for the holonomy along $c_2 \circ c_1$, respectively c^{-1}.

We also define an equivalence relation $\sim$ on $\mathcal{C}$ by $c \sim c'$ iff they only differ by a reparametrization and a finite number of retracings (a segment $s \circ s^{-1}$ of a curve as illustrated in figure 3 is called a retracing). The equivalence classes $p = [c]_\sim$ are called paths. We denote the set of paths by $\mathcal{P} := \mathcal{C}/\sim$.

The reason for this terminology is that holonomies depend only on the path equivalence class, $A(c) = A(c')$ whenever $c \sim c'$. That retracings do not affect the holonomy follows immediately from (3.4), invariance under reparametrizations is clear from the defining differential equation. It is therefore natural to define $A(p) := A(c)$ for an arbitrary representative of the path p.

In very much the same way as one can write the interaction picture propagator (time evolution operator) of ordinary quantum theory as a time-ordered exponential (see e.g. [30]), we can write the formal solution to the differential equation (3.1) as a path-ordered exponential, meaning that the terms are ordered in such a way that the curve parameters are ordered from left to right, beginning with the smallest one,

$$\begin{aligned} A(c) = \mathcal{P}\exp\left(\int_c A\right) &:= \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^1 dt_1 \ldots \int_0^1 dt_n \mathcal{P}\left(A(t_1)\ldots A(t_n)\right) \\ &= \sum_{n=0}^{\infty} \int_0^1 dt_1 \int_{t_1}^1 dt_2 \ldots \int_{t_{n-1}}^1 dt_n A(t_1)\ldots A(t_n). \end{aligned} \tag{3.5}$$

The equality of the two integral expressions can be seen by observing that the integrand

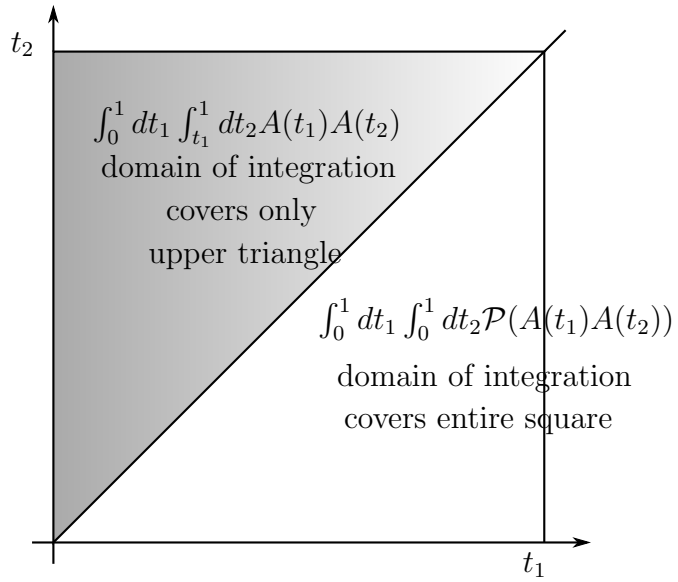

FIGURE 2. Illustration of the $n = 2$-term in equation (3.5).

of the latter is already path ordered in its domain of integration and the path ordered integrand of the first integral is "symmetric" on the cube $[0,1]^n$ (see also figure 2, which illustrates the summand with $n = 2$); to verify that the last expression indeed solves (3.1), replace the lower bound of the t_1-integral by t to obtain $h_c(t)$. We see that the holonomy $A(c)$ is (almost) the exponential of $\int_c A$. This fits of course perfectly with our task of constructing a background-independent theory, since the one-form A_a^i can be integrated over one-dimensional submanifolds without using a background metric.

Our main objection against the naive smearing was the complicated behavior under gauge transformations. Let us therefore check, if the situation is improved with the use of holonomies. We denote the holonomy of the gauge-transformed connection ${}^gA$ by ${}^gh$ and suppress the label c of the curve. ${}^gh$ satisfies the differential equation

$$ {}^g\dot{h} + {}^gA\,{}^gh = 0, \quad {}^gh(0) = \mathbb{1}. $$

Inserting the expression for the gauge-transformed connection (2.15) we have with $dg(\dot{x}) = \dot{x}(g) = \dot{g}$

$$ {}^g\dot{h} + \left(g^{-1}\dot{g} + g^{-1}Ag\right){}^gh = 0. $$

We multiply by g,

$$ g\,{}^g\dot{h} + \dot{g}\,{}^gh + Ag\,{}^gh = 0 $$

and use the Leibniz rule to arrive at

$$ (g\,{}^g\dot{h}) + A\,(g\,{}^gh) = 0. $$

$g\,{}^gh$ therefore satisfies the differential equation of the holonomy for A (but with a different initial condition). Uniqueness of solutions implies

$$ g\,{}^gh = hC $$

with a constant C which is determined to be $C = g(b(c))$ by the initial condition for ${}^g h$. Since the holonomy is the value of ${}^g h$ at $t = 1$ we arrive at

$$ {}^g h_c = g^{-1}(f(c)) h_c g(b(c)). \tag{3.6} $$

In addition to this simple behavior under gauge transformations holonomies also transform nicely under spatial diffeomorphisms φ. Since one-forms transform with the pull-back of φ we have for the coordinates

$$ A_i \rightsquigarrow (\varphi^* A)_i = A_j(\varphi(x)) \partial_i \varphi^j(x). $$

We therefore have

$$ \begin{aligned} \int_c A &= \int_0^1 A_i(x(t)) \dot{x}^i(t) dt \quad \rightsquigarrow \\ \int_c \varphi^* A &= \int_0^1 A_j(\varphi(x(t))) \partial_i \varphi^j(x(t)) \dot{x}^i(t) dt \\ &= \int_0^1 A_j((\varphi \circ x)(t)) \frac{d}{dt} (\varphi \circ x)(t) dt. \end{aligned} $$

By formula (3.5) spatial diffeomorphisms act on holonomies by simply transforming the curve,

$$ h_c(A) \rightsquigarrow h_{\varphi \circ c}(A). \tag{3.7} $$

1.2. Momentum variables. Let us now come to the conjugate variables $\widetilde{E}^a_i$. Since the connection variables have been smeared in one dimension, we will have to smear them in at least two dimensions to arrive at non-distributional Poisson brackets. Since we are not allowed to use any background structure we want to convert the electric fields into two-forms. We can do so by first adding the 3-form indices, $\widetilde{E}^a_i \underline{\delta}_{bcd} = E^a_{ibcd}$. A two-form can now be obtained by contracting the vector index with the first form-index. As this operation looks like a trace, it is natural to introduce the notation $\big(\mathrm{Tr}(\widetilde{E}^a_i)\big)_{cd} := E^a_{iacd}$.[3] This object can now be naturally integrated over two-dimensional surfaces, giving the electric fluxes

$$ E_j(S) := \int_S \mathrm{Tr}(\widetilde{E}^a_i). \tag{3.8} $$

We can also smear the fluxes with a Lie algebra valued function f^i on the surface S,

$$ E_f(S) := \int_S \mathrm{Tr}(\widetilde{E}^a_i) f^i. \tag{3.9} $$

The transformation of the fluxes under spatial diffeomorphisms can be found in the same way as the transformation of the holonomies and is again given by moving the surface S with φ. When smeared with f^i we also have to pull back this function with φ^{-1},

$$ E_f(S) \rightsquigarrow E_{(\varphi^{-1})^* f}(\varphi \circ S). \tag{3.10} $$

[3] The trace here is not to be confused with a trace of a matrix. We will therefore denote it with capital letters, whereas real matrix traces will be denoted in lower case.

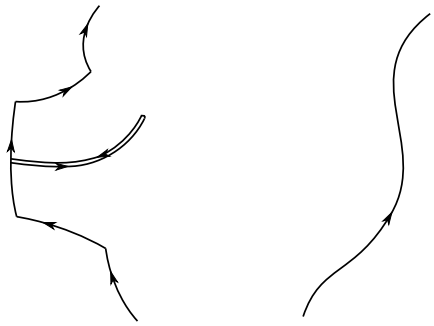

FIGURE 3. A piecewise analytic curve with a retracing on the left, contrasted with a differentiable edge on the right.

For gauge-transformations we first consider a one-parameter family of connections A_t. From ${}^g(A_t) = g^{-1}dg + g^{-1}A_t g$ we obtain the transformation

$$\dot{A} \rightsquigarrow g^{-1}\dot{A}g \tag{3.11}$$

of the velocities. Since in $\widetilde{E}^a_i \dot{A}^i_a = \mathrm{tr}(\widetilde{E}^a \dot{A}_a)$ a covector acts on a tangent vector to the configuration space producing a scalar object, this combination should be gauge invariant. By cyclicity of the trace in $\mathfrak{su}(2)$ we therefore get ${}^g\widetilde{E}^a = g^{-1}\widetilde{E}^a g$.[4]

Now the transformation of the fluxes is almost obvious (the action of g and g^{-1} can again be moved to the smearing function f by cyclicity of the trace in $\mathfrak{su}(2)$),

$${}^gE_f(S) = E_{gfg^{-1}}(S). \tag{3.12}$$

1.3. Preferred Poisson subalgebra. As our Poisson subalgebra of elementary functions that are to be quantized we take the natural construction for cotangent bundles outlined in section 1 of chapter 2. As already mentioned we do not take all smooth functions of the holonomies, but only so-called cylindrical ones. In order to define them we introduce the notion of a graph. We call a path which has an entire (i.e. not only piecewise) analytic representative an *edge*. A *graph* is then defined as a finite collection of edges that are independent in the sense that if they intersect, they only do so in their endpoints. We denote the set of edges by $E(\gamma)$ and the set of vertices (endpoints of edges) by $V(\gamma)$. Note that every finite set of paths can be made into a graph by breaking all paths that are either only piecewise analytic or that do intersect into pieces. Defining

$$p_\gamma\colon \begin{cases} \mathcal{A} \to SU(2)^{|E(\gamma)|} \\ A \mapsto (A(e))_{e\in E(\gamma)}, \end{cases}$$

[4]A conceptually cleaner way to find the transformation of the densitized triads is as follows. (3.11) defines an action Φ of the gauge group on the tangent space to $\mathcal{A}$. The (adjoint) action Φ^* on an element $\widetilde{E}^a$ of the dual space is then defined via $\langle \Phi^*(g)\widetilde{E}^a, \dot{A}_a\rangle := \langle \widetilde{E}^a, \Phi(g^{-1})\dot{A}_a\rangle$. Note that in order to get a representation we have to take g^{-1} on the RHS instead of g, so that we have $\langle \Phi^*(h)\Phi^*(g)\widetilde{E}^a, \dot{A}_a\rangle = \langle \widetilde{E}^a, \Phi(g^{-1})\Phi(h^{-1})\dot{A}_a\rangle = \langle \Phi^*(hg), \dot{A}_a\rangle$.

we call a function $f\colon \mathcal{A} \to \mathbb{C}$ cylindrical over the graph γ, provided that there exists $f_\gamma\colon SU(2)^{|E(\gamma)|} \to \mathbb{C}$ such that $f = f_\gamma \circ p_\gamma$. We will denote the set of smooth[5] cylindrical functions over γ by $\mathrm{Cyl}^\infty_\gamma$ and the set of functions that are cylindrical over some graph γ by Cyl^∞, in other words $\mathrm{Cyl}^\infty := \cup_\gamma \mathrm{Cyl}^\infty_\gamma$.

Our preferred Poisson subalgebra then consists of cylindrical functions f and fluxes $E_n(S)$,

$$\mathcal{A} := \{f + E_n(S) \mid f \in \mathrm{Cyl}^\infty, n\colon S \to \mathfrak{su}(2)\}\,. \tag{3.13}$$

To see the connection to the finite dimensional expression $v^i p_i$, note that the role of v^i is played by the constant vector field $\int_S d^2\sigma n^i(x(\sigma))\underset{\sim}{\delta}_{abc}\widetilde{\delta}(x - x(\sigma))$ which does not depend on the configuration variables.

In order for this set of functions to be appropriate for quantization, according to step 1 of the quantization program we have to show two criteria. Firstly, we have to be able to express all functions on phase space in terms of the functions in $\mathcal{A}$. We start by noting that we can recover the holonomy along an edge e from its matrix elements, which are special cylindrical functions. It then remains to show that the A^i_a can be recovered from the holonomies and the $\widetilde{E}^a_i$ from the fluxes. It is easily seen by using a sequence of curves (resp. surfaces) that shrink down to a point that we recover the values of the components of A^i_a (resp. $\widetilde{E}^a_i$) at that point. The detailed calculation does not provide too much insight and we refer the interested reader to [40].

Secondly, the most important point is that the Poisson bracket of a holonomy and a flux again gives a cylindrical function. We write A, B for matrix indices in the defining representation of $SU(2)$ and see from

$$\{f, E_i(S)\} = \sum_{e\in\gamma} \frac{\partial f_\gamma}{\partial h_e{}^A{}_B} \left\{h_e{}^A{}_B, E_i(S)\right\}$$

that we only need to calculate the bracket $\left\{h_e{}^A{}_B, E_i(S)\right\}$. To find the variation of the holonomy we vary the defining differential equation (3.1) and get the inhomogeneous linear differential equation

$$\delta\dot{h}_e + \delta A h_e + A\delta h_e = 0$$

for δh_e. The homogeneous equation is again the differential equation of the holonomy, its general solution is

$$\delta h_{e,\mathrm{hom}} = h_e C$$

with a constant C. The solution of the inhomogeneous equation can be found by variation of constants, $C = C(s)$. This leads to

$$h_e\dot{C} + \delta A h_e = 0,$$

[5]Here smooth means that the function f_γ is smooth w.r.t. the standard differentiable structure on $SU(2)^{|E(\gamma)|}$.

which is solved by

$$C(s) = C(0) - \int_0^s ds' h_e^{-1}(s')\delta A(s')h_e(s').$$

From the initial condition $\delta h_e(0) = 0$ we find $C(0) = 0$ and arrive at

$$\delta h_e(s) = -h_e(s)\int_0^s ds' h_e^{-1}(s')\delta A(s')h_e(s').$$

Denoting by $h_e(t,s)$ the holonomy along the portion of the edge e that starts at parameter value s and ends at t we can write $h_e(s)h_e^{-1}(s') = h_e(s,0)h_e(0,s') = h_e(s,s')$ by properties (3.4) of the holonomy. Evaluating at $s = 1$ we finally have

$$\delta h_e = -\int_0^1 ds h_{e_s}\delta A(s)h_{e^s} = -\int_x\int_0^1 ds h_{e_s}\dot{x}^a(s)\tau_i\widetilde{\delta}^{(3)}(x - x(s))\delta A_a^i(x)h_{e^s},$$

where e^s is the part of the edge e that starts at parameter value 0 and goes to s, whereas e_s is the part that starts from s and goes to 1.

The variation of the flux is easier to find. We parameterize the surface S by σ_1 and σ_2 and have

$$\begin{aligned}\delta E_i(S) &= \int \delta\widetilde{E}_i^a \underset{\sim}{\delta}_{abc} = \int d^2\sigma\delta\widetilde{E}_i^a(x(\sigma))\underset{\sim}{\delta}_{abc}\frac{\partial x^b}{\partial\sigma_1}\frac{\partial x^c}{\partial\sigma_2}\\ &= \int_x\int d^2\sigma\underset{\sim}{\delta}_{abc}\frac{\partial x^b}{\partial\sigma_1}\frac{\partial x^c}{\partial\sigma_2}\widetilde{\delta}^{(3)}(x - x(\sigma))\delta\widetilde{E}_i^a(x).\end{aligned}$$

We use again our standard formula (C.8) for the Poisson brackets and remember the discussion of the convolution of δ-distributions from the end of 4.2 in chapter 2 to calculate the desired bracket,

$$\{h_e, E_i(S)\} = -\frac{\beta}{2}\int ds\int d^2\sigma\widetilde{\delta}^{(3)}(x(s) - x(\sigma))\frac{\partial x^a}{\partial s}\frac{\partial x^b}{\partial\sigma_1}\frac{\partial x^c}{\partial\sigma_2}\underset{\sim}{\delta}_{abc}h_{e_s}\tau_i h_{e^s}.$$

Combining $\widetilde{\delta}^{(3)}(x(s) - x(\sigma))$ and $\underset{\sim}{\delta}_{abc}$ to obtain density weight zero we can introduce coordinates in $\widetilde{\delta}^{(3)}(x(s) - x(\sigma))\underset{\sim}{\delta}_{abc} = \delta^{(3)}(x(s) - x(\sigma))\varepsilon_{ijk}$. The factor $\varepsilon_{ijk}\frac{\partial x^i}{\partial s}\frac{\partial x^j}{\partial\sigma_1}\frac{\partial x^k}{\partial\sigma_2}$ obviously vanishes if $e \subset S$ since then the tangent vectors are linearly dependent. If they are linearly independent this factor is exactly the Jacobian from the change of coordinates $(s, \sigma_1, \sigma_2) \to (x_1, x_2, x_3)$ (only the sign is missing since we need the absolute value of the Jacobian). We therefore have

$$\{h_e, E_i(S)\} = -\frac{\beta}{2}\int d^3x\delta^{(3)}(x(s) - x(\sigma))\sigma(S,e)h_{e_s}\tau_i h_{e^s} = -\frac{\beta}{2}\sigma(S,e)h_{e_{\bar{s}}}\tau_i h_{e^{\bar{s}}}, \quad (3.14)$$

where $\bar{s}$ is the curve parameter where c intersects S and we introduced the sign function

$$\sigma(S,e) := \begin{cases} 1 & \text{if } ds\wedge d^2\sigma \text{ and } d^3x \text{ have the same orientation}\\ -1 & \text{if they have opposite orientations}\\ 0 & \text{if } e\subset S \text{ or } e\cap S = \emptyset.\end{cases}$$

We can write our final formula for $\{f, E_i(S)\}$ in a more compact form by noting that if f is cylindrical over a graph, then it is also cylindrical over a graph that is adapted to

S in the sense that every edge that intersects S starts at S (just divide edges into two parts and reverse the orientation of one the two). For edges with $b(e) \in S$ formula (3.14) acquires an extra factor $\frac{1}{2}$ because of $\int_0^1 \delta(s)ds = \frac{1}{2}$.[6] We then define the right-invariant vector-fields R^j on $SU(2)$ by their actions on functions $f\colon SU(2) \to \mathbb{R}$, ($L_g h := gh$ denotes left translation on $SU(2)$)

$$(R^j f) := \frac{d}{dt}\bigg|_{t=0} L^*_{e^{t\tau_j}} f.$$

In particular we have $R^j(h) := R^j h = \frac{d}{dt}\big|_{t=0} e^{t\tau_j} h = \tau_j h$. Defining the action of R^j_e on cylindrical functions by letting R^j act just on the holonomy along the edge e,

$$(R^j_e f_\gamma)(\{h_{e'}\}_{e'\in E(\gamma)}) = \frac{d}{dt}\bigg|_{t=0} f_\gamma(\{h_{e'}\}_{e'\neq e}, e^{t\tau_j} h_e),$$

we then arrive at

$$\{f, E_i(S)\} = -\frac{\beta}{4}\sum_{e\in\gamma} \sigma(S,e) \frac{\partial f_\gamma}{\partial h_e{}^A{}_B} (\tau_i h_e)^A{}_B = -\frac{\beta}{4}\sum_{e\in\gamma} \sigma(S,e) R^i_e f_\gamma. \tag{3.15}$$

2. Representations

One of the main strengths of LQG is a mathematically rigorous implementation of the constraints. After all, the main reason why one was stuck with the ADM formulation for so long is that operators could only be defined formally. We therefore think that it is worth treating the construction of the representation space in some detail.

Our kinematical Hilbert space will be of the form $\mathcal{H}_{\text{kin}} = L_2(\overline{\mathcal{A}}, d\mu_0)$. We therefore start by defining the configuration space $\overline{\mathcal{A}}$. Since measures are hard to find directly, it will be convenient to use the Riesz representation theorem D.4, which allows to define a measure by a corresponding linear functional on the space $C(\overline{\mathcal{A}})$. In order for this theorem to be applicable we need to make $\overline{\mathcal{A}}$ a compact Hausdorff space. This will be the content of subsection 2.1. We will then construct an orthonormal basis and use it to define the measure μ_0. Finally we construct a representation of the holonomy-flux algebra $\mathcal{A}$ on $\mathcal{H}_{\text{kin}}$.

2.1. Topology on $\overline{\mathcal{A}}$. Originally the space $\overline{\mathcal{A}}$ was constructed as the spectrum of the C^*-algebra of cylindrical functions. Fortunately there is an equivalent construction based on a projective limit, which is probably easier to understand if one is not already familiar with the theory of C^*-algebras. It is based on holonomies along curves which is motivated by the simple behavior of the objects on gauge-transformations and spatial diffeomorphisms, see (3.6) and (3.7). We follow the excellent presentation given in [40] and begin with the following

Definition 3.2. A *category* $\mathcal{K}$ consists of

- A set $\text{ob}(\mathcal{K})$ whose elements we call objects.

[6] This can heuristically be seen from $\int_0^1 \delta(s)ds = \int_{-1}^1 \Theta(s)\Theta'(s)ds = \frac{1}{2}\int_{-1}^1 (\Theta^2)'ds = \frac{1}{2}(\Theta^2\big|_1 - \Theta^2\big|_{-1}) = \frac{1}{2}$. A precise argument can be given using Colombeau theory.

- For every pair (x, y) of objects there is a set $\hom(x, y)$ of morphisms or arrows. Morphisms are often denoted by $f\colon x \to y$.
- A composition operation between morphisms,

$$\circ\colon \begin{cases} \hom(x,y) \times \hom(y,z) \to \hom(x,z) \\ (f,g) \mapsto g \circ f. \end{cases}$$

We require that the composition is associative, $h \circ (g \circ f) = (h \circ g) \circ f$, and that for every object x we have an identity element $\mathrm{id}_x \in \hom(x, x)$, satisfying $\mathrm{id}_x f = f$ for all $f \in \hom(y, x)$ and $f \circ \mathrm{id}_x = f$ for all $f \in \hom(x, y)$.

We further define a *subcategory* $\mathcal{K}'$ of $\mathcal{K}$ to be a category with $\mathrm{ob}(\mathcal{K}') \subset \mathrm{ob}(\mathcal{K})$, $\hom'(x, y) \subset \hom(x, y)$ and the same composition operation as in $\mathcal{K}$.

A morphism $f \in \hom(x, y)$ is an *isomorphism* provided that it is invertible, i.e. there exists $g \in \hom(y, x)$ such that $f \circ g = \mathrm{id}_y$ and $g \circ f = \mathrm{id}_x$.

We define a *groupoid* to be a category in which every morphism is invertible, and a *group* to be a groupoid with just one object. This definition is consistent with the conventional definition of a group. If G is a group in the conventional sense, we get a groupoid with just one abstract object, let us call it "$\cdot$". The morphisms then consist of the group elements, $\hom(\cdot, \cdot) = G$ and composition is just the group operation.

Finally, groupoid-homomorphisms are maps $F\colon \mathcal{K}_1 \to \mathcal{K}_2$ between two groupoids that preserve the structure,[7] We will only consider groupoid-homomorphisms where $\mathcal{K}_2$ is a group. For those the conditions are $F(f \circ g) = F(f) \circ F(g)$, and $F(g^{-1}) = (F(g))^{-1}$.

After this long list of definitions it is probably a good idea to develop some feeling for this new concept by verifying that some categories satisfy all required axioms. An easy example is the category whose objects are all sets, the morphisms are the functions between two sets and composition is just composition of functions. Another category consists of all topological spaces, the morphisms are then all homeomorphisms and composition is again composition of functions. Similarly one can define the categories of groups and homomorphisms or vector spaces and linear mappings.

To see the connection of this new terminology to our task of constructing the space $\overline{\mathcal{A}}$, we observe that the set $\mathcal{P}$ of paths in Σ forms a groupoid, as is easy to see. To be precise, the objects are the points of Σ and the morphisms between two points are the paths between those points. Now we already saw that holonomies are actually defined on paths rather than curves. From equations (3.4) we have

$$A(p \circ p') = A(p)A(p') \text{ and } A(p^{-1}) = (A(p))^{-1}, \tag{3.16}$$

telling us that every connection A is a groupoid homomorphism between the groupoids $\mathcal{P}$ and $SU(2)$. Now the definition of the space $\overline{\mathcal{A}}$ is almost obvious.

[7] In mathematical terms these maps are called covariant *functors*.

DEFINITION 3.3. The configuration space $\overline{\mathcal{A}}$ consists of all groupoid-homomorphisms between the groupoids $\mathcal{P}$ and $SU(2)$, in other words all functions satisfying (3.16).

It can be shown that $\overline{\mathcal{A}}$ contains elements that are in a certain sense non-continuous, which means that the space $\mathcal{A}$ of all connections is a *proper* subset of $\overline{\mathcal{A}}$. One can also say that the space $\overline{\mathcal{A}}$ is a distributional extension of the space $\mathcal{A}$.

For the construction of the compact topology we need another mathematical concept. Recall that a *partial order* $\prec$ on a set $\mathcal{L}$ is a relation (a subset of $\mathcal{L} \times \mathcal{L}$) that is reflexive ($l \prec l$), antisymmetric ($l \prec l'$ and $l' \prec l$ imply $l = l'$) and transitive ($l \prec l'$ and $l' \prec l''$ imply $l \prec l''$). If the relation instead is only reflexive and transitive, but also satisfies the condition that for all l, l' there is an l'' such that $l \prec l''$ and $l' \prec l''$, then $(\mathcal{L}, \prec)$ is called a *directed* set.

DEFINITION 3.4. A projective family $(X_l, p_l^{l'})$ is a family of sets $(X_l)_{l\in\mathcal{L}}$, where the index l runs in a partially ordered and directed set $\mathcal{L}$. Whenever $l \prec l'$ we have a surjective "projection" map $p_l^{l'} : X_{l'} \to X_l$. Two successive projections have to satisfy

$$p_l^{l'} \circ p_{l'}^{l''} = p_l^{l''}, \quad \forall l \prec l' \prec l''. \tag{3.17}$$

The projective limit $\overline{X}$ of $(X_l, p_l^{l'})$ is defined by

$$\overline{X} := \left\{ (x_l) \in \prod_{l\in\mathcal{L}} X_l \mid p_l^{l'}(x_{l'}) = x_l, \forall l \prec l' \right\}. \tag{3.18}$$

The motivation for considering the projective limit is as follows. The elements x_l will be identified with restrictions $A|_{l(\gamma)}$ and the projections $p_l^{l'}$ will also be restrictions. The statement that $A|_l|_{l'} = A|_{l'}$ then translates to the condition in (3.18). We call sequences $(x_l)_{l\in\mathcal{L}}$ which satisfy this condition coordinated.

We will show in a minute that $\overline{\mathcal{A}}$ can be identified with a projective limit. The idea is as follows. Since $\overline{\mathcal{A}}$ consists of mappings into the compact group $SU(2)$, it is natural to construct the projective family in such a way that the X_l are subsets of $\overline{\mathcal{A}}$. It is then possible to use the compactness of $SU(2)$ to induce a compact topology on the X_l. By Tychonov's theorem $\prod_{l\in\mathcal{L}} X_l$ is then also compact. Since closed subspaces of a compact space are themselves compact, it then remains to show that $\overline{X}$ is a closed subset of $\prod_{l\in\mathcal{L}} X_l$.

Let us now outline the details of the construction. In order to specify the projective family, we have to define a partially ordered index set $\mathcal{L}$, the sets X_l and the projections $p_l^{l'}$. To find the index set, we note that every graph γ defines a subgroupoid $l(\gamma)$ of $\mathcal{P}$ by taking as objects the vertices of the graph and as morphisms the set of edges, their inverses, and finite compositions of composable edges.

DEFINITION 3.5. The index set $\mathcal{L}$ of our projective family consists of all subgroupoids $l(\gamma)$ of $\mathcal{P}$ generated by graphs. The relation $\prec$ on $\mathcal{L}$ is given by $l(\gamma) \prec l(\gamma')$ iff $l(\gamma)$ is a subgroupoid of $l(\gamma')$.

The fact that $(\mathcal{L}, \prec)$ is a partially ordered set is almost obvious. It is also clear that we have $l(\gamma) \prec l(\gamma'')$ and $l(\gamma') \prec l(\gamma'')$ with $\gamma'' := \gamma \cup \gamma'$. But we have to be careful here, since it is not clear that $\gamma \cup \gamma'$ defines a graph. Remember that the union of graphs is to be understood in the sense that all intersecting edges have to be divided into pieces. In order to arrive at a graph again, we must make sure that $\gamma \cup \gamma'$ only consists of a finite number of edges. This is equivalent to showing that two edges can only intersect in a finite number of points (or they are analytic continuations of each other, since in that case we do not need to break them into pieces). Suppose that the edges e and e' intersect in an infinite discrete set. We can choose suitably parameterized representatives c and c' such that $c^i(t) - c'^i(t)$ is a compactly supported analytic function that vanishes on an infinite discrete set. But the following lemma then shows that this function must vanish identically, giving $e = e'$. Note that for the lemma to be applicable analyticity is crucial.

Lemma 3.6. *A compactly supported, analytic, real valued function $f\colon [0,1] \to \mathbb{R}$ that vanishes at an infinite number of points vanishes identically.*

Proof. By the Bolzano-Weierstrass theorem, the set of the zeros of the function f has an accumulation point t_0 in $[0,1]$. By analyticity of f it can be expanded in an absolutely convergent power series $f(t) = \sum_{n=0}^{\infty} d_n (t - t_0)^n$. If we can show that all the coefficients vanish, then $f = 0$ and the proof is complete. Suppose therefore that this is not the case. Then there is a smallest integer k, such that $d_k \neq 0$. We can then write $f(t) = (t - t_0)^k g(t)$ with $g(t) = \sum_{m=0}^{\infty} d_{k+m} (t - t_0)^m$. The function g is continuous at t_0 and $g(t_0) = d_k \neq 0$. Thus $g(t) \neq 0$ in a whole neighborhood of t_0, and therefore also $f(t) \neq 0$ in this neighborhood, which is a contradiction to t_0 being an accumulation point of zeros of f. □

Definition 3.7. X_l is the set of all homomorphisms from the subgroupoid $l(\gamma)$ to the group $SU(2)$.

Since every element of X_l has to satisfy equations (3.16), it is completely specified by the values it takes on the edges e of the graph. We can therefore identify X_l and $SU(2)^{|E(\gamma)|}$. Since $SU(2)$ is a compact topological group and since the finite product of compact spaces is compact, X_l carries a natural compact topology.

Definition 3.8. The projections $p_l^{l'}(x_{l'})$ are given by simply restricting the action of the homomorphism $x_{l'} \in X_{l'}$ to the elements of the subgroupoid $l(\gamma)$.

That those projections satisfy the consistency condition (3.17) is obvious. That they are also surjective is more tricky since it does *not* follow from $l(\gamma) \prec l(\gamma')$ that every edge of the graph γ is also an edge of the graph γ'. We can only state that every $e \in \gamma$ is a finite composition of the edges of γ' and their inverses,

$$e = \circ_{e' \in E(\gamma')} (e')^{s_{ee'}}. \tag{3.19}$$

Since the edges of a graph can only intersect in their endpoints, every edge e' can only appear once in this decomposition, which means $s_{ee'} \in \{0, \pm 1\}$. We have already seen that every $x_l \in X_l$ is determined by its values $x_l(e) = h_e$ on all the edges of the graph γ, and it is also clear that every possible choice of values h_e yields an element of X_l. Showing that every x_l is the restriction of an $x_{l'}$ (i.e. surjectivity) then means that for every choice of $(h_e)_{e \in E(\gamma)}$ we can find a family $(h_{e'})_{e' \in \gamma'}$, such that for all edges e we have (here we denote the product in $SU(2)$ also by $\circ$)

$$h_e = \circ_{e' \in E(\gamma')} (h_{e'})^{s_{ee'}}.$$

Now the crucial observation is that every edge of γ' that appears in the decomposition of an edge $e \in \gamma$ cannot appear in the decomposition of another edge $\widetilde{e} \in \gamma$ (here the fact that the edges of γ cannot overlap enters). For every $e \in \gamma$ we can therefore take any any $e'(e)$ in the decomposition of e and set $h_{e'(e)} := h_e$. All other $h_{e'}$ can be set to identity if they are not of the form $h_{e'(e)}$ with $e \in \gamma$.

For the following we will again need to introduce some mathematical machinery. A net $(x_\alpha)_{\alpha \in A}$ in a topological space X is basically the same as a sequence, with the important difference that the index set A is not $\mathbb{N}$, but instead a directed set. In mathematical terms a net therefore is a mapping $x \colon A \to X$. Convergence is defined in the same way as for sequences: A net (x_α) converges to x, denoted $\lim_{\alpha \in A} x_\alpha = x$ iff for every neighborhood U of x there is an index $\alpha(U)$ such that $x_\alpha \in U$ for all $\alpha(U) \prec \alpha$. We then have the following statement, which is not true for sequences (at least not in both directions). A function $f \colon X \to Y$ between topological spaces is continuous iff for every convergent net $(x_\alpha) \to x$ in X, the net $(f(x_\alpha))$ converges to $f(x)$.

We also need to know how the product topology on a direct product of topological spaces is defined.

Definition 3.9. The product topology on the product $\prod_{l \in \mathcal{L}} X_l$ of topological spaces X_l is the weakest topology such that all projections $p_k \colon \prod_{l \in \mathcal{L}} X_l \to X_k$; $(x_l) \mapsto x_k$ are continuous.

It can be shown that a net $x^\alpha = (x_l^\alpha)$ converges to $x = (x_l)$ in the product topology, iff for every $l \in \mathcal{L}$, (x_l^α) converges to (x_l) (this is another example of an assertion that is not true for sequences). We then have the following powerful result,

Theorem 3.10 (Tychonov). *The space $\prod_{l \in \mathcal{L}} X_l$ is compact in the product topology iff every X_l is compact.*

For a finite product the theorem can actually be shown by elementary means. The amazing statement is that it also holds for arbitrary index sets $\mathcal{L}$. The proof uses the axiom of choice and can be found for example in [40].

The theorem of Tychonov guarantees that the infinite product $\prod_{l \in \mathcal{L}} X_l$ is compact. Being a subspace of this space, we can equip the projective limit $\overline{X}$ with the subspace

topology. If we can manage to show that $\overline{X}$ is closed, it is automatically compact since every closed subspace of a compact space is compact.

PROPOSITION 3.11. *The set $\overline{X}$ of coordinated sequences is a closed subspace of $\prod_{l\in\mathcal{L}} X_l$.*

PROOF. We start by showing that the projections $p_l^{l'}$ are continuous using the above mentioned characterization of continuity by the convergence of nets. With the decomposition (3.19) and the identification of X_l with $SU(2)^{|E(\gamma)|}$ the projection maps are given by

$$p_l^{l'}\left((h_{e'})_{e'\in E(\gamma')}\right) = \Big(\prod_{e'\in E(\gamma')} (h_{e'})^{s_{ee'}}\Big)_{e\in E(\gamma)},$$

where the $h_{e'}$ are the ones from the proof of surjectivity. Assume now that the net $\left((h_{e'}^{\alpha})_{e'\in E(\gamma')}\right)_{\alpha\in A}$ converges in $SU(2)^{|E(\gamma')|}$ to $(h_{e'})_{e'\in E(\gamma')}$. By the definition of the product topology every component $(h_{e'}^{\alpha})_{\alpha\in A}$ converges to $h_{e'}$. Since $SU(2)$ is a Lie group, multiplication and inversion are continuous, hence $\left(\prod_{e'\in E(\gamma')}(h_{e'}^{\alpha})^{s_{ee'}}\right)_{\alpha\in A}$ converges to $\prod_{e'\in E(\gamma')}(h_{e'})^{s_{ee'}}$ for every e. Finally we use again the definition of the product topology to conclude that $\left(\left(\prod_{e'\in E(\gamma')}(h_{e'}^{\alpha})^{s_{ee'}}\right)_{e\in E(\gamma)}\right)_{\alpha\in A}$ converges to $\left(\prod_{e'\in E(\gamma')}(h_{e'})^{s_{ee'}}\right)_{e\in E(\gamma)}$, which concludes the proof of continuity of the projections.

To show that $\overline{X}$ is closed we use the fact that a subset Y of a topological space X is closed iff every net (x_α) that converges in X but has all $x_\alpha \in Y$ has its limit lying in the subspace Y. Let therefore $((x_l^\alpha)_{l\in\mathcal{L}})$ be a net in $\overline{X}$ and suppose that it converges in $\prod_{l\in\mathcal{L}} X_l$ to $(x_l)_{l\in\mathcal{L}}$. We have

$$p_l^{l'}(x_{l'}) = p_l^{l'}(\lim_\alpha x_{l'}^\alpha) = \lim_\alpha p_l^{l'}(x_{l'}^\alpha) = \lim_\alpha x_l^\alpha = x_l,$$

because of the definition of the product topology, continuity of the projections $p_l^{l'}$, $(x_l^\alpha)_{l\in\mathcal{L}}$ lying in $\overline{X}$ and again the definition of the product topology. □

Note that the space $\overline{X}$ is also Hausdorff. This is clear from the fact that $SU(2)$ is Hausdorff (as is every manifold) and the general statement that (arbitrary) direct products and subspaces of Hausdorff spaces are again Hausdorff.

It remains to identify $\overline{X}$ with the space $\overline{\mathcal{A}}$ of generalized connections that we originally wanted to equip with a compact and Hausdorff topology.

PROPOSITION 3.12. *The function*

$$\Phi\colon \begin{cases} \overline{\mathcal{A}} \to \overline{X} \\ A \mapsto (A|_l)_{l\in\mathcal{L}}, \end{cases}$$

that maps every generalized connection A to the family of restrictions to subgroupoids generated by a graph, is a bijection.

PROOF. That $\Phi(A)$ is an element of $\overline{X}$ is obvious. To show injectivity, suppose that $\Phi(A) = \Phi(A')$. Then $A|_l = A'|_l$ for *every* subgroupoid l. Therefore $A(p) = A'(p)$ for every path p, and $A = A'$.

To show surjectivity, let $(x_l)_{l\in\mathcal{L}} \in \overline{X}$. We consider the function $A(p) := x_{l_p}(p)$ where l_p is an arbitrary subgroupoid containing the path p. Once we have shown that the so-defined A is well defined and indeed an element of $\overline{\mathcal{A}}$, it is clear that $\Phi(A) = (x_l)_{l\in\mathcal{L}}$. The latter assertion is clear from the fact that x_l is a groupoid-homomorphisms from l to $SU(2)$. That A is well-defined means that it is independent of the subgroupoid used in its definition. Let l' therefore be another subgroupoid containing the path p. Then we find l'' such that $l \prec l''$ and $l' \prec l''$ (because $\mathcal{L}$ is directed). Since $(x_l)_{l\in\mathcal{L}}$ is an element of $\overline{X}$, both x_l and $x_{l'}$ are the restriction of $x_{l''}$. Therefore $x_l(p) = x_{l''}(p) = x_{l'}(p)$. □

2.2. Spin networks and the Ashtekar measure. As the title suggests, some knowledge of measure theory is probably helpful to understand the contents of this section. However, the intuitive idea of a measure that a typical undergraduate physicist has, should be sufficient here. The reader who still feels uneasy is encouraged to look at the first chapters of Rudin's book [34]. Some basic material on representations of Lie groups is provided in appendix D.

As already advertised, we will use the Riesz representation theorem to construct the measure μ_0. We therefore have to specify a linear functional Λ_0 on $C(\overline{\mathcal{A}})$. A convenient way to do this is to define the values of Λ_0 on the following set of functions, whose linear span will be seen to be dense in $C(\overline{\mathcal{A}})$.

DEFINITION 3.13.

(1) A spin network $s = \big(\gamma, \{j_e\}_{e\in E(\gamma)}, \{m_e\}_{e\in E(\gamma)}, \{n_e\}_{e\in E(\gamma)}\big)$ (in the following abbreviated by SNW) consists of a graph γ, in which every edge e is assigned a half-integral spin quantum number $j_e \in \{\frac{1}{2}, 1, \frac{3}{2}, \dots\}$ and two magnetic quantum numbers $m_e, n_e \in \{-j_e, -j_e+1, \dots, j_e\}$.

(2) The spin-network function T_s of the SNW s (abbreviated by SNWF) is defined by

$$T_s : \begin{cases} \overline{\mathcal{A}} \to \mathbb{C} \\ A \mapsto \prod_{e\in E(\gamma)} \sqrt{2j_e+1}\,(\rho_{j_e}(A(e)))_{m_e n_e}\,, \end{cases} \tag{3.20}$$

where $\rho_j(h)$ denotes the matrix of the group element h in the spin-j representation and $(\rho_j(h))_{mn}$ are the elements of this matrix.

An example of a simple SNW (without magnetic quantum numbers) is given in figure 4.

LEMMA 3.14. *The SNWF are linearly independent. Their finite linear combinations span* $\{T_s\}$ *form a dense subspace of* $C(\overline{\mathcal{A}})$.

PROOF. To see that the SNWF are linearly independent, consider a linear combination of the zero vector $\sum_i \lambda_i T_{s_i} = 0$. We can multiply by $\overline{T}_{s_j}$ and integrate over all edges using the Haar measure on $SU(2)$ to obtain $\sum_i \lambda_i \int \overline{T}_{s_j} T_{s_i} (d\mu_H)^N = 0$. The integral

$\int \overline{T}_s T_{s'} (d\mu_H)^N$ consists of factors $\int d\mu_H = 1$ (normalization of μ_H; integrals corresponding to edges not contained in γ and γ'), factors $\int d\mu_H(A(e)) \left(\rho_{j_e}(A(e))\right)_{m_e n_e} = 0$ (by the Peter Weyl theorem after inserting $1 = (\rho_0(A(e)))_{11}$; integrals corresponding to edges contained in only one of the graphs). Finally for edges contained in both graphs we have a factor $\int d\mu_H(A(e)) \left(\rho_{j_e}(A(e))\right)_{m_e n_e} \left(\rho_{j'_e}(A(e))\right)_{m'_e n'_e} = (2j_e+1)^{-1} \delta_{j_e j'_e} \delta_{m_e m'_e} \delta_{n_e n'_e}$ again by the Peter Weyl theorem. We thus have $\int \overline{T}_s T_{s'} (d\mu_H)^N = \delta_{ss'}$ and conclude that in our linear combination of the zero vector $\lambda_j = 0$. Since j was arbitrary, our linear combination is trivial which proves linear independence of the SNWF.

That span $\{T_s\}$ is dense in $C(\overline{\mathcal{A}})$ can be shown by checking that the hypotheses of the Stone-Weierstrass theorem D.3 are fulfilled. We have already seen that $\overline{\mathcal{A}}$ is a compact Hausdorff space. That our subspace is nowhere vanishing and separates the points of $\mathcal{A}$ is straightforward to show. Slightly more interesting is the fact that it is closed under complex conjugation. By unitarity of the representation matrices we have

$$\overline{T_s(A)} = \prod_{e\in\gamma} \sqrt{(2j_e+1)}\, \overline{\left(\rho_{j_e}(A(e))\right)}_{m_e n_e} = \prod_{e\in\gamma} \sqrt{(2j_e+1)}\, \left(\rho_{j_e}(A(e^{-1}))\right)_{n_e m_e},$$

which is the SNWF of the SNW with the orientation of all edges reversed and the magnetic quantum numbers exchanged. Since span $\{T_s\}$ is obviously a linear subspace it remains to show that the product of two SNWF is again a SNWF. If the graphs are disjoint it is clear that this product is the SNWF of the union of the graphs (where all edges keep their quantum numbers). In the case that some edges intersect we have to divide those into two parts, $e = e_2 \circ e_1$. The corresponding factor in the SNWF is then $\left(\rho_j(A(e))\right)_{mn} = \sum_p \left(\rho_j(A(e_2))\right)_{mp} \left(\rho_j(A(e_1))\right)_{pn}$, which leads again to a linear combination of SNWF. The most complicated case arises when an edge is contained in both SNW-s. The product of the SNWF then contains the factor $\left(\rho_{j_1}(A(e))\right)_{m_1 n_1} \left(\rho_{j_2}(A(e))\right)_{m_2 n_2}$. This is a matrix element in the tensor product $j_1 \otimes j_2$, which is a linear combination of matrix elements of the irreducible representations $|j_1 - j_2|, \ldots, j_1 + j_2$, according to the Clebsh Gordan theorem (see section 1.1 of appendix D). □

THEOREM 3.15.

(1) The functional Λ_0, whose values on the SNWF are given by

$$\Lambda_0(T_s) := \begin{cases} 1 & s = (\emptyset, 0, 0, 0) \\ 0 & \textit{otherwise}, \end{cases} \tag{3.21}$$

extends to a positive and continuous linear functional on $C(\overline{\mathcal{A}})$.

(2) For cylindrical functions over the graph γ the value of Λ_0 is given by

$$\Lambda_0(f) = \int_{SU(2)^{|E(\gamma)|}} \prod_{e\in E(\gamma)} d\mu_H(h_e) f_\gamma(\{h_e\}_{e\in E(\gamma)}). \tag{3.22}$$

PROOF. Because of the linear independence of the SNWF, Λ_0 is a well defined linear functional on span $\{T_s\}$. The linear span of the SNWF carries a norm (the supremum

norm), hence continuity of Λ_0 can be checked using sequences. Let therefore a sequence of linear combinations of SNWF be given that converges to zero, $\lim_{n\to\infty}\sum_s \lambda_{s,n}T_s = 0$. Then, by linear independence of the T_s, we have $\lim_{n\to\infty}\lambda_{s,n} = 0$ for all s and hence also $\lim_{n\to\infty}\Lambda_0(\sum_s \lambda_{s,n}T_s) = \lim_{n\to\infty}\lambda_{s_0,n} = 0$, where s_0 denotes the empty spin network, which proves continuity of Λ_0. Since span $\{T_s\}$ is a dense subspace of $C(\overline{\mathcal{A}})$, Λ_0 extends uniquely to a continuous linear functional on $C(\overline{\mathcal{A}})$.

Before we check positivity we verify formula (3.22). For a SNWF (which is obviously cylindrical over the graph of its SNW) we have

$$\int_{SU(2)^{|E(\gamma)|}} \prod_{e\in E(\gamma)} d\mu_H(h_e)(T_s)_\gamma(\{h_e\}_{e\in E(\gamma)})$$

$$= \prod_{e\in E(\gamma)} \int_{SU(2)} d\mu_H(h_e)\sqrt{(2j_e+1)}\,(\rho_{j_e}(h_e))_{m_e n_e} = \begin{cases} 1 & s = (\emptyset, 0, 0, 0) \\ 0 & \text{otherwise,} \end{cases}$$

by the Peter Weyl theorem (after inserting a factor 1 like in the proof of linear independence of the SNWF). By linearity of Λ_0 and the integral, the formula extends to span $\{T_s\}$. We already showed that span $\{T_s\}$ is dense in $C(\overline{\mathcal{A}})$. By the same reasoning one can show that the linear span of the SNWF whose graph is a subgraph of γ is dense in $C(X_{l(\gamma)})$. Since the linear span of those SNWF is contained in Cyl_γ, which is in turn contained in $C(X_{l(\gamma)})$, every function $f \in \mathrm{Cyl}_\gamma$ is a (uniform) limit of such SNWF. Therefore, by continuity of Λ_0 and the limit $f_n \to f$ being uniform (so that we can exchange limit and integral)

$$\Lambda_0(f) = \lim_{n\to\infty}\Lambda_0(f_n) = \lim_{n\to\infty}\int_{SU(2)^{|E(\gamma)|}} \prod_{e\in E(\gamma)} d\mu_H(h_e)(f_n)_\gamma(\{h_e\}_{e\in E(\gamma)})$$

$$= \int_{SU(2)^{|E(\gamma)|}} \prod_{e\in E(\gamma)} d\mu_H(h_e) f_\gamma(\{h_e\}_{e\in E(\gamma)}).$$

From expression (3.22) for cylindrical functions and the Peter Weyl theorem now easily follows

$$\Lambda_0(\overline{T}_s T_{s'}) = \delta_{ss'}, \tag{3.23}$$

which proves positivity of Λ_0. □

We have now verified that Λ_0 satisfies all the assumptions of the Riesz representation theorem. It therefore defines a probability measure μ_0 via

$$\Lambda_0(f) = \int_{\overline{\mathcal{A}}} d\mu_0(A) f(A).$$

Definition 3.16. The kinematical Hilbert space is defined as $\mathcal{H}_{\text{kin}} := L_2(\overline{\mathcal{A}}, d\mu_0)$. It has the SNWF T_s as orthonormal basis (see (3.23)) and is therefore non-separable.

2.3. Representation of the holonomy-flux algebra $\mathcal{A}$. We are now ready to define a representation of the holonomy-flux algebra (3.13) on $\mathcal{H}_{\text{kin}}$. We do this in the standard way already outlined in section 1 of chapter 2. Cylindrical functions $f \in \mathrm{Cyl}^\infty(\overline{\mathcal{A}})$

act as multiplication operators on $g \in \mathcal{H}_{\text{kin}}$,

$$(\widehat{f}g)(A) := (fg)(A), \tag{3.24}$$

whereas the fluxes $E_j(S)$ act as derivative operators and are therefore only defined on the dense subspace of once differentiable cylindrical functions,

$$(\widehat{E_n(S)}g)(A) := i\frac{\beta}{2}\left\{g, E_n(S)\right\}(A). \tag{3.25}$$

We now have to verify that this representation meets our criteria discussed in chapter 2, section 1. That it satisfies the commutation relations (2.3) is clear, since it is just the standard representation (2.4). It remains to verify the star relations and that all operators are self-adjoint. For the multiplication operators this is obvious, since the adjoint is just complex conjugation. The corresponding result for the derivative operators is slightly more difficult to show. We refer the interested reader to [40].

3. Implementation of the kinematical constraints

The implementation of the kinematical constraints (i.e. the Gauß and diffeomorphism constraint) is beautifully explained in Rovelli's book [32]. We follow the line of his treatment and start by implementing the Gauß constraint. The resulting space of solutions will be denoted by $\mathcal{H}_0$. The diffeomorphism constraint will then be defined on this space, the resulting Hilbert space of diffeomorphism invariant states will be called $\mathcal{H}_{\text{diff}}$. The last step will then be to implement the Hamiltonian constraint on $\mathcal{H}_{\text{diff}}$ — we postpone this issue until section 5 because the implementation will use the volume operator. We therefore have the following sequence, where each step consist in the solution of a constraint,[8]

$$\mathcal{H}_{\text{kin}} \xrightarrow{SU(2)} \mathcal{H}_0 \xrightarrow{\text{Diff}^*} \mathcal{H}_{\text{diff}} \xrightarrow{\text{Hamiltonian}} \mathcal{H}.$$

3.1. Gauß constraint. Remembering that the Gauß constraint just implements invariance under local gauge transformations (see section 3.1), and looking at the behavior (3.6) of holonomies under these transformations, it is natural to consider the following representation of the gauge group $SU(2)$. We define the action $\widehat{U}$ of $SU(2)$ on $\mathcal{H}_{\text{kin}}$ densely on cylindrical functions via

$$(\widehat{U}(g)f)_\gamma(A) := f_\gamma(\{(g(f(e)))^{-1}A(e)g(b(e))\}_{e\in E(\gamma)}).$$

It is easy to see that this representation of the gauge group is unitary,

$$\begin{aligned}\left(\widehat{U}(g)f, \widehat{U}(g)f\right) &= \int_{G^{|E(\gamma)|}} \prod_{e\in E(\gamma)} (d\mu_H(h_e)) \left|f_\gamma\left(\left\{(g(f(e)))^{-1}h_e g(b(e))\right\}_{e\in E(\gamma)}\right)\right|^2 \\ &= \int_{G^{|E(\gamma)|}} \prod_{e\in E(\gamma)} (d\mu_H(h_e)) \left|f_\gamma\left(\{h_e\}_{e\in E(\gamma)}\right)\right|^2 = (f,f)\,,\end{aligned}$$

[8]Actually it turns out that the Hamiltonian constraint can not be defined as an operator on $\mathcal{H}_{\text{diff}}$. We will discuss this issue in section 5.2.

due to left and right invariance of the Haar measure. This again contributes to the appeal of the chosen representation, since the natural actions on Hilbert spaces to consider are those which preserve the inner product. $\widehat{U}(g)$ is thus in particular an isometric operator on the cylindrical functions and can be uniquely extended to an isometric operator on $\mathcal{H}_{\text{kin}}$.

The space of solutions to the Gauß constraint consists of elements $f \in \mathcal{H}_{\text{kin}}$ which are invariant under gauge transformations, $\widehat{U}(g)f = f$. To describe this space we need intertwiners, which are introduced in appendix D. To each vertex v of a SNW we assign an intertwiner i_v between the representations j_e of all edges that begin or end at that vertex. As explained in the aforementioned appendix, we lower every index corresponding to an edge that *ends* at that vertex, so that this index is now invariant under the *dual* representation. From the transformation of the SNWF,

$$\widehat{U}(g) \prod_{e\in E(\gamma)} \rho_{j_e}(A(e))^{\alpha_e}{}_{\beta_e} = \prod_{e\in E(\gamma)} \rho_{j_e}((g(f(e)))^{-1})^{\alpha_e}{}_{\alpha'_e} \rho_{j_e}(A(e))^{\alpha'_e}{}_{\beta'_e} \rho_{j_e}(g(b(e)))^{\beta'_e}{}_{\beta_e}$$

and from the invariance of the intertwiners (D.3), it is then clear that the following linear combinations of SNWF form a basis of the space of solutions to the Gauß constraint (we abbreviate the number of vertices of γ by V and the number of edges by E),

$$\begin{aligned} |S\rangle = \sum_{\alpha_l,\beta_l} & i_{v_1}{}^{\beta_1\ldots\beta_{n_1}}{}_{\alpha_1\ldots\alpha_{m_1}} i_{v_2}{}^{\beta_{n_1+1}\ldots\beta_{n_2}}{}_{\alpha_{m_1+1}\ldots\alpha_{m_2}} \\ & \ldots i_{v_V}{}^{\beta_{n_{V-1}+1}\ldots\beta_{n_E}}{}_{\alpha_{m_{V-1}+1}\ldots\alpha_{m_E}} \prod_l \rho_{j_l}(A(e_l))^{\alpha_l}{}_{\beta_l}. \end{aligned} \tag{3.26}$$

In order to better explain this formula we note again that every index α_l is contracted with an index on an intertwiner at the vertex where e_l ends, and every index β_l is contracted with an index of an intertwiner at a vertex where e_l starts. Note that these states are labeled by a graph γ, together with a spin quantum number j_e for each edge (and no magnetic quantum numbers, since these indices are contracted in (3.26)).[9] We call graphs labeled in this way gauge invariant spin networks S and the solutions $|S\rangle$ gauge invariant spin network functions. It is easy to see that the gauge invariant SNWF form an orthonormal basis of the space $\mathcal{H}_0$ of solutions to the Gauß constraint. Note however that this basis is not unique, since a different choice of intertwiners gives a different basis.

Example 3.17. As an example consider the SNW of figure 4 with two vertices and three edges, each starting at v_1 and ending at v_2. The edges carry the representations $j_1 = 1$, $j_2 = j_3 = \frac{1}{2}$. An intertwiner between the three representations is an invariant vector in the tensor product $j_1 \otimes j_2 \otimes j_3$. It therefore has to satisfy

$$R(U)^i{}_j U^A{}_C U^B{}_D i^{j,CD} = i^{i,AB},$$

[9]We do not consider the intertwiners as labels here since a different choice of intertwiners yields a state that is linear dependent.

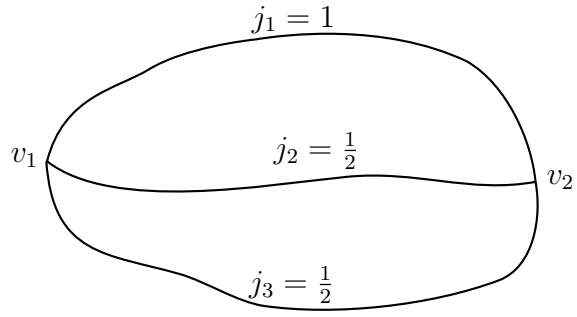

FIGURE 4. A simple SNW.

where the index i of the adjoint representation corresponds to components w.r.t. the basis in the space of symmetric 2-spinors, as explained in footnote 1 of appendix D. It can be shown that this equation is satisfied by

$$i^{i,AB} := \frac{1}{\sqrt{3}}\sigma^{i,AB} := \frac{1}{\sqrt{3}}\sigma_i{}^A{}_C\varepsilon^{BC},$$

where $\sigma_i{}^A{}_B$ are the Pauli matrices. The gauge-invariant spin network state of this simple spin network is therefore

$$|S(A)\rangle = \frac{1}{3}\sigma_{i,AB}(R(A(e_1)))^i{}_j(A(e_2))^A{}_C(A(e_3))^B{}_D\sigma^{j,CD}.$$

3.2. Diffeomorphism constraint. Similarly to the Gauß constraint we implement the diffeomorphism constraint by requiring invariance under diffeomorphisms. Actually we can not consider the group of *all* diffeomorphisms, but rather need to consider the group $\mathrm{Diff}^*(\Sigma)$, which consists of all continuous bijections $\phi\colon \Sigma \to \Sigma$ such that ϕ and ϕ^{-1} are analytic everywhere except at a finite number of isolated points. The reason is that diffeomorphisms that are only smooth and not analytic could destroy analyticity of the curves of a given graph. On the other hand, transformed curves may fail to be analytic at a finite number of points — but they remain piecewise analytic.

From the transformation of the holonomies (3.7), it is natural to consider the representation $\widehat{V}$ defined densely on cylindrical functions via (note that contrary to the action on holonomies (3.7), in order to get a representation on a space of functions of holonomies we have to act with ϕ^{-1})

$$(\widehat{V}(\phi)f)_\gamma(A) := f_\gamma(\{A(\phi^{-1}(e))\}_{e\in E(\gamma)}) = f_{\phi^{-1}\circ\gamma}(\{A(e)\}_{e\in E(\phi^{-1}\circ\gamma)}).$$

Since the inner product between cylindrical functions does not explicitly depend on the graph, unitarity of the representation is obvious from

$$\left(\widehat{V}(\phi), \widehat{V}(\phi)\right) = \int d\mu_H^{|\gamma|}\overline{f}_{\phi^{-1}\circ\gamma}f_{\phi^{-1}\circ\gamma} = (f, f)\,.$$

It turns out that there are no solutions to the diffeomorphism constraint in $\mathcal{H}_{\text{kin}}$. To see this, suppose that $\sum c_s T_s$ is gauge-invariant. Since the T_s form an orthonormal basis, every coefficient $c_{s'}$ corresponding to a SNW with graph $\phi(\gamma)$ has to be equal to the coefficient corresponding to the SNW with graph γ and the same labeling of the edges. It follows that there have to be uncountably many non-vanishing coefficients — clearly a contradiction. We therefore have to use the method outlined briefly in point 3 of the quantization program in section 1 of chapter 2 and look for solutions in the dual space of the cylindrical functions.

Example 3.18. As a warm up we consider the following easy examples of single particles obeying ordinary quantum mechanics (no gravitation).

(1) Take a single particle that is allowed to move only on the x-axis and search for translation invariant wave functions. The translation operators T_a, $a \in \mathbb{R}$ are defined by $(T_a f)(x) = f(x+a)$. Translation invariant states have to satisfy $T_a f = f$, for all $a \in \mathbb{R}$ and therefore have to be constant functions, $f(x) = f_0$. Unfortunately, constant functions are clearly not in the Hilbert space $L^2(\mathbb{R})$. However, we can consider the Gelfand triple

$$C_0^\infty(\mathbb{R}) \hookrightarrow L^2(\mathbb{R}) \hookrightarrow C_0^\infty(\mathbb{R})',$$

where $C_0^\infty(\mathbb{R})$ are the smooth functions of compact support and $C_0^\infty(\mathbb{R})'$ their dual space (space of continuous linear functionals or distributions). The constant function f_0 can then be identified with the distribution $\langle f_0, g\rangle := f_0 \int g(x)dx$. This allows us to systematically construct solutions by introducing a rigging map η, which maps functions in the dense subspace $C_0^\infty(\mathbb{R})$ to distributional solutions. In our case, $\eta(f) = \int da T_a \overline{f}(x) = \int da \overline{f}(x+a) = \int da \overline{f}(a) = \overline{f}_0$, where the constant value $\overline{f}_0$ is identified with the corresponding distribution — in other words $\eta(f)(g) = \overline{f}_0 \int g dx = \overline{f}_0 g_0$. It is also easy to define an inner product on the translation invariant distributions via $\langle \eta(f), \eta(g)\rangle := \eta(f)(g) = \overline{f}_0 g_0$. (Observe that the naive definition of the inner product by interpreting the distributions as constant functions and (ab)using the L^2 inner product does not work since $\int dx f_0 \overline{g}_0 = \infty$.)

The distributional nature of the translation invariant wave functions is probably more apparent in the momentum representation. Here T_a acts via multiplication, $(T_a f)(p) = e^{ipa} f(p)$. Obviously all translation invariant wave functions have the form $f(p) = c\delta(p)$ with a constant c. The rigging map is defined again by averaging over the action of the group, $\eta(f) = \int da e^{-ipa} \overline{f}(p) = \delta(p)\overline{f}(p) = \delta(p)\overline{f}(0)$. We therefore have $\eta(f)(g) = \overline{f}(0)g(0)$, which is the same result as before since the value of $f(p)$ at $p = 0$ equals the constant f_0 from above, as is clear from $f(p) = \int f(x)e^{-ipx}dx$. Note that in the momentum representation the naive definition of the inner product would lead to multiplication of δ-distributions!

(2) A slightly more interesting example is that of a particle that is allowed to move in a plane. The configuration space is $T^*\mathbb{R}^2$ with canonical Poisson brackets between position and momentum variables. The particle shall be constrained by $C(x^1, x^2, p_1, p_2) := p_1 = 0$. We consider again the momentum representation with representation space $L^2(\mathbb{R}^2)$ where position operators act by differentiation, $\widehat{x}^i f(p_1, p_2) = -i\frac{\partial}{\partial p_i} f(p_1, p_2)$ and momentum operators act by multiplication, $\widehat{p}_i f(p_1, p_2) = p_i f(p_1, p_2)$. All operators are defined on the dense subspace $\mathcal{S}(\mathbb{R}^2)$ of smooth functions of rapid decrease. We therefore have the Gelfand triple

$$\mathcal{S}(\mathbb{R}^2) \hookrightarrow L_2(\mathbb{R}^2) \hookrightarrow \mathcal{S}'(\mathbb{R}^2),$$

where $\mathcal{S}'(\mathbb{R}^2)$ is the space of tempered distributions, i.e. the space of continuous linear functionals on $\mathcal{S}(\mathbb{R}^2)$. It is obvious that for every function $f \in \mathcal{S}(\mathbb{R}^2)$ the distribution

$$\eta(f) := \delta(p_1) f(p_1, p_2) = \delta(p_1) f(0, p_2)$$

is a solution to the constraint. The inner product on the space of solutions is

$$\begin{aligned}\langle \eta(f), \eta(g) \rangle &:= \eta(f)(g) = (\delta(p_1) f(p_1, p_2))(g(p_1, p_2)) \\ &= \int dp_2 \overline{f}(0, p_2) g(0, p_2).\end{aligned}$$

Let us now return to the issue of solving the diffeomorphism constraint. We have already solved the Gauß constraint and denoted the space of solutions by $\mathcal{H}_0$. This space is a subspace of $\mathcal{H}_{\text{kin}}$ and therefore inherits a Hilbert space structure. We restrict the action of the representation $\widehat{V}$ to the dense subset $\mathcal{S}_0 \subset \mathcal{H}_0$, which consists of all *finite* linear combinations of gauge invariant SNWF. We therefore have the Gelfand triple[10]

$$\mathcal{S}_0 \hookrightarrow \mathcal{H}_0 \hookrightarrow \mathcal{S}_0^*.$$

Since $(\widehat{V}'(\phi)\Phi)(f) = \Phi(\widehat{V}^\dagger(\phi) f) = \Phi(\widehat{V}(\phi^{-1}) f)$, our space of solutions consists of all linear functionals $\Phi \in \mathcal{S}_0^*$ that satisfy

$$\Phi(\widehat{V}(\phi) f) = \Phi(f), \quad \forall \phi \in \text{Diff}^*(\Sigma), f \in \mathcal{S}_0. \tag{3.27}$$

We now define the rigging map $\eta \colon \mathcal{S}_0 \to \mathcal{S}_0^*$ by

$$\eta(f)(g) = \sum_{f' \in \text{Orb}(f)} \langle f', g \rangle,$$

where the orbit of $f \in \mathcal{S}_0$ is defined by

$$\text{Orb}(f) := \left\{ f' \in \mathcal{S}_0 \mid \exists \phi \in \text{Diff}^*(\Sigma), \text{ such that } f' = \widehat{V}(\phi) f \right\}.$$

Note here that being in the same orbit obviously defines an equivalence relation on $\mathcal{S}_0$. First of all, the map η is well defined, since the sum is always finite. This is due to the

[10]To be precise, $\mathcal{S}_0^*$ is the space of *all* linear functionals on $\mathcal{S}_0$ (not just the continuous ones). It is naturally equipped with the weak-$*$ topology. In other words, a net (Φ_α) converges to Φ iff the net $\Phi_\alpha(f)$ of complex numbers converges to $\Phi(f)$ for all $f \in \mathcal{D}_{\text{kin}}$.

orthogonality of the gauge invariant SNWF, which form a basis of $\mathcal{S}_0$. If a diffeomorphism changes the graph of the SNW, the corresponding summand is therefore zero. If it does not change the graph (it might for example just reverse the orientation of some of its edges), the corresponding summand does not vanish, but there are only a finite amount of functions in the orbit of f that are reached by such diffeomorphisms.[11]

That $\eta(f)$ is invariant under diffeomorphisms can be seen from

$$\begin{aligned}\eta(f)(\widehat{V}(\phi)g) &= \sum_{f'\in\mathrm{Orb}(f)} \left\langle f', \widehat{V}(\phi)g\right\rangle = \sum_{f'\in\mathrm{Orb}(f)} \left\langle \widehat{V}(\phi)\underbrace{\widehat{V}(\phi^{-1})f'}_{=:f''}, \widehat{V}(\phi)g\right\rangle \\ &= \sum_{f''\in\mathrm{Orb}(f)} \langle f'', g\rangle = \eta(f)(g).\end{aligned}$$

In fact, it is also easily shown that the image $\eta(\mathcal{S}_0)$ is dense in the space of all solutions to the diffeomorphism constraint. Any $\Phi \in \mathcal{H}_{\mathrm{diff}}$ is completely determined by its action on SNWF and has to satisfy

$$\Phi(T_s) = \Phi(\widehat{V}(\phi)T_s), \quad \forall\phi \in \mathrm{Diff}^*.$$

In other words, Φ has to take the same values on all $T_{s'}$ in the orbit of a given T_s, but other than that its action on the SNWF is arbitrary. If Φ is only nonzero on a finite number of orbits, it is clear that we can choose $f \in \mathcal{S}_0$ with $\Phi = \eta(f)$ (Every $f \in S_0$ is a finite linear combination $\sum_s c_s T_s$. If $\Phi(T_{s'}) \neq 0$ then only one coefficient c_s in the orbit of $T_{s'}$ has to be chosen appropriately.) Now if Φ is nonzero on infinitely many orbits, we can choose a sequence of f_n, such that $\eta(f_n)(T_s) \to \Phi(T_s), \forall T_s$ (this works because we will see below that there are only a countable number of orbits). By the choice of topology on $\mathcal{S}_0^*$ detailed in footnote 10 we then also have $\eta(f_n) \to \Phi$.

From our introductory examples it is also clear how to define the inner product on the space of solutions,

$$\langle \eta(f), \eta(g)\rangle = \eta(f)(g).$$

Finally, lets address a criticism that is often brought forward against LQG. We have seen before that the SNWF basis of $\mathcal{H}_{\mathrm{kin}}$ is uncountable, the Hilbert space therefore not separable. Many physicists therefore claim that LQG quantizes the "wrong" Hilbert space, since every physically reasonable Hilbert space has to be separable (although there is no convincing argument for this statement). Nevertheless, the diffeomorphism constraint reduces the large Hilbert space $\mathcal{H}_{\mathrm{kin}}$ to a separable Hilbert space $\mathcal{H}_{\mathrm{diff}}$ (observe also that this reduction takes place before any dynamics enters the stage, since dynamics is completely contained in the Hamiltonian constraint).

Since obviously $\eta(f) = \eta(g)$ if f and g are in the same orbit, we can identify the elements of $\eta(\mathcal{S}_0)$ with equivalence classes in $\mathcal{S}_0$. An equivalence class of graphs under diffeomorphisms is called a *knot*. We therefore see that the basis states in $\mathcal{H}_{\mathrm{diff}}$ are labeled

[11]There is of course an uncountable number of diffeomorphisms that yield a non-vanishing summand. But the sum goes over the functions in the orbit of f and not over the diffeomorphisms.

by knots. It is a key result of knot theory that the set of knots is countable, therefore $\mathcal{H}_{\text{diff}}$ is separable.[12] For a slightly more detailed discussion of the separability of $\mathcal{H}_{\text{diff}}$, and also for an intuitive treatment of the connection to the choice of Diff* over the ordinary diffeomorphism group, refer to [32].

4. Kinematical geometrical operators

We discuss in this section operators corresponding to the area of a coordinate surface and the volume of a coordinate region. An operator for the length of a curve can also be constructed, but this is much more difficult to achieve (one has to use the quantization technique of Poisson brackets with the volume operator, which is outlined for the Hamiltonian constraint below). Those operators will be gauge-invariant. But since they do not commute with the diffeomorphism constraint, the question of their physical significance arises. It can be shown that one can for example define coordinate surfaces with respect to matter, without changing the spectrum, and thereby obtain a diffeomorphism invariant operator. For a sketch of the proof see [40]. Unfortunately those operators still do not commute with the Hamiltonian constraint and it is not yet known how to deal with this issue.

4.1. Area operator. In order to find a quantization, we first need to express the area of a coordinate surface in terms of our basic operators $E_i(S)$ and $A(e)$. For this purpose we consider the following definition. Let $\omega_{ab,cd}$ be a tensor on Σ, antisymmetric in its first, as well as in its last two indices and symmetric under the exchange of ab and cd. Then we define

$$\int_S \sqrt{\omega_{ab,cd}} := \int_U \sqrt{\omega_{ij,kl}\frac{\partial x^i}{\partial \sigma_1}\frac{\partial x^j}{\partial \sigma_2}\frac{\partial x^k}{\partial \sigma_1}\frac{\partial x^l}{\partial \sigma_2}}d^2\sigma,$$

where $S\colon U \subset \mathbb{R}^2 \to \Sigma$ is a surface embedded in Σ and parameterized by σ_1, σ_2. In particular we can consider the tensor

$$G_{ab,cd} := \delta^{mn}_{ab} g_{mc} g_{nd},$$

for which

$$\begin{aligned}\int_S \sqrt{G_{ab,cd}} &= \int_S \sqrt{g_{ac}g_{bd} - g_{cb}g_{da}} = \int_U \sqrt{(g_{ik}g_{jl} - g_{kj}g_{li})\frac{\partial x^i}{\partial \sigma_1}\frac{\partial x^j}{\partial \sigma_2}\frac{\partial x^k}{\partial \sigma_1}\frac{\partial x^l}{\partial \sigma_2}}d^2\sigma \\ &= \int_U \sqrt{\det S^* g}\, d^2\sigma = A(S)\end{aligned}$$

is the area of the surface S. Up to now we could also have considered p-dimensional submanifolds and just found an alternative way of rewriting its p-dimensional volume. For a surface S in the three-dimensional manifold Σ however, we can rewrite $G_{ab,cd}$ in

[12] Of course there is the additional complication coming from the different coloring of the graphs. But again there is only a finite amount of different colorings for a given graph.

terms of the quantities $\widetilde{E}^a_i \widetilde{\delta}_{abc}$ that appear in $E_i(S)$,

$$\begin{aligned}\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \widetilde{E}^m_i \underset{\sim}{\delta}_{mnl} &= E^a_i \varepsilon_{abc} E^m_i \varepsilon_{mnl} = h^{am} \varepsilon_{abc} \varepsilon_{mnl} \\ &= \varepsilon^m{}_{bc} \varepsilon_{mnl} = h_{bq} h_{ck} \underbrace{\varepsilon^{mqk} \varepsilon_{mnl}}_{=\delta^{qk}_{nl}} = G_{bc,nl}.\end{aligned}$$

We conclude that

$$A(S) = \int_S \sqrt{\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \widetilde{E}^m_i \underset{\sim}{\delta}_{mnl}}.$$

Still, this is not yet the expression we were looking for, because the two-forms $\widetilde{E}^a_i \underset{\sim}{\delta}_{abc}$ in the integrand are not integrated. However, we can write the integral as a Riemann sum with the Riemann decomposition $\mathcal{R} = ((\Delta S_I)_{I\in\mathcal{I}}, (x_I)_{I\in\mathcal{I}})$ (so that $S = \cup_I \Delta S_I$ and $x_I \in \Delta S_I$),

$$\begin{aligned}\int_S \sqrt{\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \widetilde{E}^m_i \underset{\sim}{\delta}_{mnl}} &= \int_U \sqrt{\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \widetilde{E}^m_i \underset{\sim}{\delta}_{mnl} \frac{\partial x^b}{\partial \sigma_1} \frac{\partial x^c}{\partial \sigma_2} \frac{\partial x^n}{\partial \sigma_1} \frac{\partial x^l}{\partial \sigma_2}} d^2\sigma \\ &= \lim_{|\mathcal{R}|\to 0} \sum_I \sqrt{\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \widetilde{E}^m_i \underset{\sim}{\delta}_{mnl} \frac{\partial x^b}{\partial \sigma_1} \frac{\partial x^c}{\partial \sigma_2} \frac{\partial x^n}{\partial \sigma_1} \frac{\partial x^l}{\partial \sigma_2}}\Bigg|_{x_I} A(\Delta S_I) \\ &= \lim_{|\mathcal{R}|\to 0} \sum_I \sqrt{\left(\widetilde{E}^a_i \underset{\sim}{\delta}_{abc} \frac{\partial x^b}{\partial \sigma_1} \frac{\partial x^c}{\partial \sigma_2}\right)\Bigg|_{x_I} A(\Delta S_I) \left(\widetilde{E}^m_i \underset{\sim}{\delta}_{mnl} \frac{\partial x^n}{\partial \sigma_1} \frac{\partial x^l}{\partial \sigma_2}\right)\Bigg|_{x_I} A(\Delta S_I)} \\ &= \lim_{|\mathcal{R}|\to 0} \sum_I \sqrt{E_i(\Delta S_I) E_i(\Delta S_I)},\end{aligned}$$

where we have used the mean value theorem for the integral $E_i(\Delta S_I)$. Let us again summarize our classical expression for the area of a coordinate surface in terms of elementary operators,

$$A(S) = \lim_{|\mathcal{R}|\to 0} \sum_I \sqrt{E_i(\Delta S_I) E_i(\Delta S_I)}. \tag{3.28}$$

We can now promote the $E_i(\Delta S_I)$ to operators and compute the action of the area operator on SNWF. Note that this makes sense, since $E_i(\Delta S_I) E_i(\Delta S_I)$ is a sum of squares of self-adjoint operators. The spectrum is therefore positive and we can use the spectral theorem to define the square root. From the definition (3.25) of $E_i(\Delta S_I)$ as a derivative operator and from (3.15) we see that the action of $E_i(\Delta S_I)$ just inserts a matrix ${}^{(j)}\tau_i$ (a generator of $\mathfrak{su}(2)$ in the spin-j representation) times $-i\frac{\beta}{2}$ at the intersection point of ΔS_I with an edge of the SNW. The second operator $E_i(\Delta S_I)$ inserts another matrix $-i\frac{\beta}{2}\,{}^{(j)}\tau_i$. Now if the surface intersects the SNW in more than one point, ${}^{(j)}\tau_i$-matrices at different points get contracted and we do not get a gauge invariant state. Fortunately due to the limit of the Riemann sum the surface shrinks down until there is only one intersection point left (such as in the partition of figure 5). For this intersection point we have a factor $-\frac{\beta^2}{4}\,{}^{(j)}\tau_i\,{}^{(j)}\tau_i$. Since $\tau_i = -\frac{i}{2}\sigma_i$, we know that the Lie algebra generators are proportional to the operators of angular momentum in a given direction, $\tau_i = -iJ_i$.

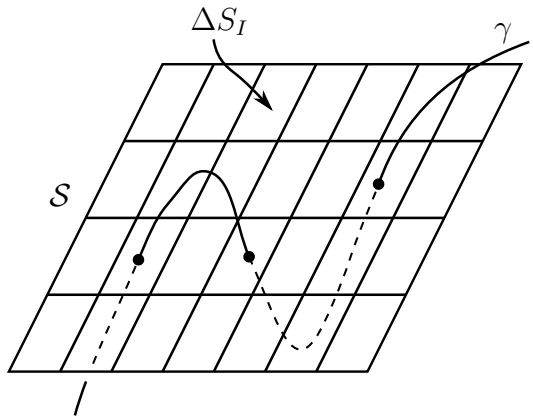

FIGURE 5. A partition of $\mathcal{S}$, fine enough so that every every ΔS_I only contains one intersection point with the graph γ.

Therefore $\tau_i\tau_i = -J^2$ and we have ${}^{(j)}\tau_i\,{}^{(j)}\tau_i = -j(j+1)\mathbb{1}$.[13] We also see that further refinement of the Riemann decomposition of S is irrelevant, as soon as every ΔS_I cuts only one edge of the SNW, because if it does not cut any edge, the contribution vanishes. We can therefore write the action of the area operator on a SNW-state as

$$A(S)\,|S\rangle = \frac{\beta}{2}\sum_{P\in(S\cap\gamma)}\sqrt{j_P(j_P+1)}\,|S\rangle\,, \tag{3.29}$$

where the sum runs over all intersection points of S and the SNW γ of $|S\rangle$ and j_P is the spin of the intersecting edge. We see that the spectrum of the area operator is labeled by multiplets $j = (j_1, \ldots, j_n)$ with n arbitrary and the eigenvalues are given by[14]

$$A_j = \frac{\beta}{2}\sum_i \sqrt{j(j+1)}. \tag{3.30}$$

The physical conclusion is that, since the spectrum of the area operator is discrete, the area of a physical surface is quantized. Restoring the physical units that we have set equal to unity, the area operator is given by $l_P^2 = \frac{16\pi G\hbar}{c^3}$ times (3.29). The minimum possible area of a surface arises if S just cuts one edge with $j = \frac{1}{2}$ and is given by

$$A_0 = \frac{\sqrt{3}\beta}{4}l_P^2, \tag{3.31}$$

[13]This can also be seen directly by noting that ${}^{(j)}\tau_i\,{}^{(j)}\tau_i$ commutes with all the ${}^{(j)}\tau_i$-matrices, hence it must be proportional to the identity operator (this is an easy corollary to Schur's lemma D.1 for the case that the two representations coincide). The proportionality constant can be calculated by evaluating the trace.

[14]Actually we have just computed what is called the main sequence of the spectrum, since we left out the action of $A(S)$ on SNWF with vertices on the surface S. However, Rovelli argues that the physically interesting part of the spectrum is given by the main sequence [32].

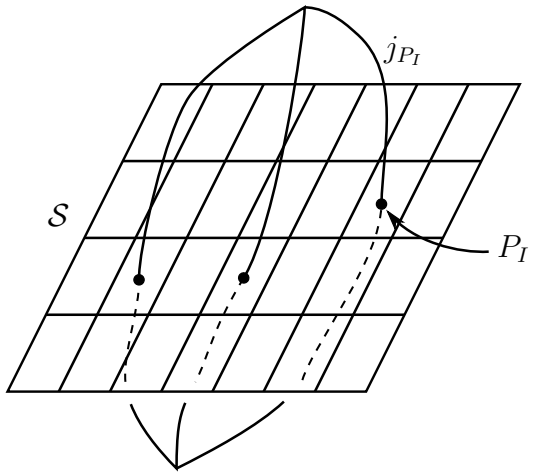

FIGURE 6. Action of the area operator on a simple SNW at the intersection points P_I.

which is of the order of the Planck area, at least as long as β is of the order of unity. This is a strong result of LQG, since it has long been expected that a theory of quantum gravity should lead to discreteness of physical space at the order of the Planck length [29]. In contrast to other theories, we have not postulated this discreteness — instead it is a direct consequence of our quantization of general relativity.

4.2. Volume operator. The operator corresponding to the volume of a given three-dimensional coordinate region $\mathcal{R} \subset \Sigma$ is not only interesting for its direct physical interpretation, but it also plays a key role in the definition of the Hamiltonian constraint (see below).

From the classical formula (2.27) we immediately have the classical expression of the volume,

$$V(\mathcal{R}) = \int_{\mathcal{R}} \omega_h = \int_{\mathcal{R}} \sqrt{\frac{1}{3!} \left| \underset{\sim}{\delta}_{abc} \varepsilon^{ijk} \widetilde{E}^a_i \widetilde{E}^b_j \widetilde{E}^c_k \right|}. \tag{3.32}$$

Unfortunately writing this expression as a Riemann sum does not directly yield the fluxes, as was the case for the area operator. One therefore has to use a regularization procedure, the details of which are quite involved. We will therefore be content with listing some of the essential features of the resulting operator.

First of all, let us mention that there exist two versions of the volume operator, corresponding to different types of regularization procedures. The two versions were developed by Rovelli and Smolin [33], resp. Ashtekar and Lewandowski [6]. Although the resulting operators are different, their qualitative features are the same. Both versions can be show to be self-adjoint, non-negative operators. As was the case for the area operator, the spectrum is discrete and there is a fundamental volume gap.

Unfortunately there is no closed formula for the action of the volume operator on a given SNWF. Still, it can be shown that it acts only on vertices and in particular for

gauge-invariant SNWF it changes the intertwiner structure of the vertices located inside the region $\mathcal{R}$. Also the action is only non-trivial on vertices with at least four incident edges — the action on 3-valent vertices vanishes due to gauge invariance. Unfortunately the work involved in calculating the action of the volume operator on a vertex increases drastically with the amount of edges incident at this vertex.

5. Implementation of the Hamiltonian constraint

The quantization of the Hamiltonian constraint (2.22) has for a long time been the greatest obstacle to finding a quantization of general relativity. The main problem was thought to be the tremendous nonlinearity contained in the factor ω_h^{-1}. The original motivation to introduce the new variables was therefore to choose $\beta = \pm i$ for the Immirzi parameter and multiply $\widetilde{H}$ by ω_h, which leads to a polynomial expression. Unfortunately the multiplication with ω_h also changes the density weight from one to two. However, only densities of weight one can be integrated in a coordinate independent manner and have a chance of giving a diffeomorphism invariant operator.[15] An obvious solution would be to take the square-root of the resulting expression to arrive at the correct density weight of one. However, attempts in that direction were not successful — mainly due to the complications involved in taking the square-root of an operator.

Another problem with this approach is that choosing β imaginary leads to an $SL(2,\mathbb{C})$ connection as opposed to an $SU(2)$ connection. Since $SL(2,\mathbb{C})$ is not compact, all the machinery we have developed is meaningless. One is then confronted with the same problems as in the original ADM formalism — due to the lack of a mathematically rigorous implementation only formal solutions could be found. For these reasons one keeps the factor ω_h^{-1} and takes β real. The solution to the nonlinearity of the Hamilton constraint contained in the factor ω_h^{-1} was found by Thiemann in the mid 1990ies [37]. We present the basic ideas below.

5.1. Classical reformulation in terms of Poisson brackets. The crucial observation is that one can trade the factor ω_h^{-1} for a Poisson bracket with the volume operator. The idea is then to quantize Poisson brackets by $\frac{1}{i\hbar}$ times the commutator. The following lemma contains the classical reformulation of the Hamiltonian constraint in terms of variables which can be quantized unambiguously.

Lemma 3.19.

(1) For $\beta = \pm i$ the Hamiltonian constraint (2.22) *can be written as*

$$H_E(N) = \frac{2}{3\beta} \int_\Sigma N \widetilde{\delta}^{abc} F^i_{ab} \left\{ A^i_c, V \right\}, \tag{3.33}$$

where $V(x)$ is the volume of an arbitrary open neighborhood $\mathcal{R}_x$ of x.

[15] In Yang-Mills theories one uses renormalization techniques to deal with operators of density weight different from one. However, every renormalization unavoidably introduces a length scale. This procedure is therefore unacceptable for us, since we want to construct a *background independent* theory.

(2) The non-Euclidean part can be written as

$$H(N) - H_E(N) = -\frac{8\left(1+\beta^2\right)}{3\beta^3}\int_\Sigma N\widetilde{\delta}^{abc}\varepsilon^{ijk}\left\{A_a^i, \overline{K}\right\}\left\{A_b^j, \overline{K}\right\}\left\{A_c^k, V\right\}, \tag{3.34}$$

with the mean integrated curvature $\overline{K}$, defined by $\overline{K} := \int_\Sigma K_a^i \widetilde{E}_i^a$. In terms of elementary variables $\overline{K}$ reads $\overline{K} = \{H_E(1), V\}$.

The motivation for the expression of $\overline{K}$ in terms of the time-evolution of the volume operator stems from

$$\mathcal{L}_n \int \omega_h = \int \frac{1}{2}\omega_h h^{ab}\mathcal{L}_n h_{ab} = \int \omega_h K = \overline{K}$$

by (B.2) and (A.7) when we regard the derivative in direction of n^a as a time-derivative.

Proof.

(1) We start with the Poisson bracket $\{A_a^i, V\}$ and observe that

$$\delta_{\widetilde{E}} V(x) = \int_{\mathcal{R}_x} \delta\omega_h(x) = \int_\Sigma \frac{1}{2} e_a^i \mathbb{1}_{\mathcal{R}_x} \delta\widetilde{E}_i^a, \tag{3.35}$$

by equation (2.31), which gives

$$\begin{aligned}\left\{A_c^i(x), V(x)\right\} &= \frac{\beta}{2}\int_z \frac{\delta A_c^i(x)}{\delta A_b^j(z)}\frac{\delta V(x)}{\delta \widetilde{E}_j^b(z)} \\ &= \frac{\beta}{2}\int_z \frac{1}{2} e_b^j(z)\mathbb{1}_{\mathcal{R}_x}(z)\delta_c^b\delta_j^i\widetilde{\delta}(x-z) = \frac{\beta}{4}e_c^i(x).\end{aligned}$$

For the Euclidean part it remains to show the equality

$$\frac{2}{3\beta}\widetilde{\delta}^{abc}\left\{A_c^i, V\right\} = \varepsilon^{lji}\widetilde{E}_l^a\widetilde{E}_j^b\omega_h^{-1}.$$

We can prove this identity on the basis $(\widetilde{E}_i^d)_{i=1,2,3}$,

$$\frac{2}{3\beta}\widetilde{\delta}^{abc}\frac{\beta}{4}e_c^i\widetilde{E}_i^d = \frac{1}{3!}\widetilde{\delta}^{abd}\omega_h = \varepsilon^{lji}\widetilde{E}_l^a\widetilde{E}_j^b\widetilde{E}_i^d\omega_h^{-1}$$

and contract with $\underset{\sim}{\delta}_{abd}$ (since both sides are completely antisymmetric we only need to verify one component),

$$\omega_h = \varepsilon^{lji}\underset{\sim}{\delta}_{abd}\widetilde{E}_l^a\widetilde{E}_j^b\widetilde{E}_i^d\omega_h^{-1} = \omega_h$$

by equation (2.27).

(2) Again we start with the relevant Poisson bracket,

$$\begin{aligned}\left\{A_a^i(x), \overline{K}\right\} &= \frac{\beta}{2}\,\delta_{\widetilde{E}}\overline{K}\big|_{\delta\widetilde{E}_j^b(z)=\delta_a^b\delta_j^i\widetilde{\delta}(x-z)} \\ &= \frac{\beta}{2}\int_z \left(K_b^j(z)\delta\widetilde{E}_j^b(z) + \delta K_b^j(z)\widetilde{E}_j^b(z)\right)\Big|_{\delta\widetilde{E}_j^b(z)=\delta_a^b\delta_j^i\widetilde{\delta}(x-z)} = \frac{\beta}{2}K_a^i(x),\end{aligned}$$

where the term with the variation of the extrinsic curvature vanishes for similar reasons as in the calculation of the Poisson bracket in section 4.3 of chapter 2: δK_a^i

is proportional to $\delta\omega_a^i$, which is in turn given by (2.35). However, we only need the variation evaluated at $\delta\widetilde{E}_j^b(z) = \delta_a^b\delta_j^i\widetilde{\delta}(x-z)$. We are left with $\int D\widetilde{\delta}(x-z)\times$"terms that are annihilated by D". Since the action of D on a distribution is defined by letting the formally adjoint operator (which is $-D$) act on the corresponding test function (i.e. the other terms under the integral) the term vanishes.

The expression (3.34) now easily follows from the corresponding formula (3.33) for the Euclidean part.

To express the mean integrated curvature in terms of elementary variables, we use the original expression (1.11) for the Hamiltonian constraint. We have

$$\begin{aligned}\delta_A H_E(1) &= \delta_A \int \left({}^3R + K^2 - K_{ab}K^{ab}\right)\omega_h = \int \left(2K\delta K - 2K_a^b\delta K_b^a\right)\omega_h \\ &= \int \left(2K\widetilde{E}_i^a\delta K_a^i - 2K_a^b\widetilde{E}_i^a\delta K_b^i\right) = \frac{2}{\beta}\int \left(K\widetilde{E}_i^a - K_b^a\widetilde{E}_i^b\right)\delta A_a^i.\end{aligned}$$

Together with (3.35) we conclude

$$\{H_E(1), V\} = \frac{\beta}{2}\frac{1}{\beta}\int (3K - K)\,\omega_h = \overline{K}. \qquad \square$$

From part (2) of lemma 3.19 we see that the additional term for Lorentzian signature contains the Euclidean constraint via the expression for $\overline{K}$. It is therefore natural to first construct an operator corresponding to the Euclidean Hamiltonian constraint and then use this operator to quantize the full constraint. For simplicity we will only outline the basic steps for the Euclidean constraint. The construction of the operator for the full constraint is completely parallel.

The plan should now be almost obvious. We take the expression (3.33), replace $V(x)$ by the volume operator for the neighborhood $\mathcal{R}_x$ and the Poisson bracket by a commutator. Unfortunately we do not have an operator for the connection (appearing also in the curvature F_{ab}). We therefore have to apply a regularization to replace them by corresponding holonomies. For that purpose we take a series of triangulations T_ε of Σ (a decomposition of Σ into mutually disjoint tetrahedra whose union is Σ). In the limit $\varepsilon \to 0$ the fineness of the triangulation tends to zero, hence every tetrahedron will be contained in a single coordinate patch for sufficiently small ε. We can therefore perform all subsequent calculation in an adapted coordinate frame and use the following lemma.

Lemma 3.20.

(1) Let $p_{\varepsilon,x}^u$ be a path with beginning point x and tangent vector εu at x, then

$$h_{p_{\varepsilon,x}^u}(A) = \mathbb{1}_2 - \varepsilon A_a(x)u^a + \varepsilon^2\left((A_a(x)u^a)^2 - \partial_b A_a(x)u^b u^a\right) + O(\varepsilon^3). \tag{3.36}$$

(2) Let $\alpha_{\varepsilon,x}^{uv}$ be a loop starting at x with tangent vectors εu and $-\varepsilon v$ at x for parameter values 0 and 1 respectively, then

$$h_{\alpha_{\varepsilon,x}^{uv}} = \mathbb{1}_2 - \varepsilon^2 F_{ab}(x)u^a v^b + O(\varepsilon^3). \tag{3.37}$$

Proof.

(1) The first statement is just the Taylor expansion of the function $\epsilon \mapsto h_{p^u_{\varepsilon,x}}(A)$. The derivatives can be found with the help of an arbitrary path p such that $p_\varepsilon(t) = p(\varepsilon t)$.[16] We then have $h_{p_\varepsilon}(A) = h_p(\varepsilon)$, which yields the leading term from the initial condition (3.1). The differential equation (3.1) for the holonomy gives the next term,

$$\frac{d}{d\varepsilon} h_{p_\varepsilon}(A)\Big|_{\varepsilon=0} = \dot{h}_p(\varepsilon)\Big|_{\varepsilon=0} = -A_a(p(\varepsilon))\dot{p}^a(\varepsilon)h_p(\varepsilon)|_{\varepsilon=0} = -A_a(x)u^a. \tag{3.38}$$

For the ε^2-term we use that up to ε^2-corrections the path is given by $x(t) = x + \varepsilon ut$, and thus $\ddot{x}(t) = 0$. We differentiate the middle expression of (3.38) according to the Leibniz rule and have

$$\begin{aligned}\frac{d^2}{d\varepsilon^2} h_{p_\varepsilon}(A)\Big|_{\varepsilon=0} &= -\dot{A}_a(p(\varepsilon))\dot{p}^a(\varepsilon)h_p(\varepsilon) - A_a(p(\varepsilon))\dot{p}^a(\varepsilon)\dot{h}_p(\varepsilon)\Big|_{\varepsilon=0} \\ &= -\partial_b A_a(x)u^b u^a + (A_a(x)u^a)^2 .\end{aligned}$$

(2) In any coordinate neighborhood the loop can be approximated by the composition of three paths,

$$\alpha^{uv}_{\varepsilon,x} = p^{-v}_{\varepsilon,x+\varepsilon v} \circ p^{v-u}_{\varepsilon,x+\varepsilon u} \circ p^u_{\varepsilon,x} + O(\varepsilon^2).$$

This allows us to use the result from the first part of the lemma. Together with the expansion $A_a(x+\varepsilon v) = A_a(x) + \varepsilon v^b \partial_b A_a(x) + O(\varepsilon^2)$ (and similar for u) we obtain

$$\begin{aligned}h_{\alpha^{uv}_{\varepsilon,x}} = &\left(\mathbb{1} + \varepsilon\left(A_a + \varepsilon v^b \partial_b A_a\right) v^a + \frac{\varepsilon^2}{2}\left((A_a v^a)^2 - \partial_a A_b v^a v^b\right)\right) \\ &\times \Bigg(\mathbb{1} - \varepsilon\left(A_a + \varepsilon u^b \partial_b A_a\right)(v^a - u^a) \\ &\qquad + \frac{\varepsilon^2}{2}\left((A_a(v^a - u^a))^2 - \partial_a A_b (v^a - u^a)\left(v^b - u^b\right)\right)\Bigg) \\ &\times \left(\mathbb{1} - \varepsilon A_a u^a + \frac{\varepsilon^2}{2}\left((A_a u^a)^2 - \partial_a A_b u^a u^b\right)\right) + O(\varepsilon^3).\end{aligned}$$

After carefully expanding in powers of ε we see that all terms proportional to ε cancel, as well as most ε^2-terms. We remember the definition (2.14) and arrive at

$$\mathbb{1} - \frac{\varepsilon^2}{2}\left(u^b \partial_b A_a v^a - v^b \partial_b A_a u^a + A_a u^a A_b v^b - A_a v^a A_b u^b\right) = \mathbb{1} - \varepsilon^2 F_{ab} u^a v^b. \qquad \square$$

Now it is easy to find the regulated expression of the Hamiltonian constraint. We denote the tetrahedra of the triangulation by Δ. For every tetrahedron we single out one of its vertices and call it $v(\Delta)$. The three edges starting at $v(\Delta)$ will be denoted by $p_I(\Delta)$, $I = 1, 2, 3$. The triangular loop consisting of $p_I(\Delta)$, the edge $a_{IJ}(\Delta)$ connecting the other endpoints of $p_I(\Delta)$ and $p_J(\Delta)$ and the inverse edge $p_J(\Delta)$ is denoted by $\alpha_{IJ}(\Delta)$ (thus $\alpha_{IJ}(\Delta) := p_J^{-1}(\Delta) \circ a_{IJ}(\Delta) \circ p_I(\Delta)$). For an illustration of this notation see figure 7.

[16] For simplicity we abbreviate $p^u_{\varepsilon,x}$ by p_ε.

PROPOSITION 3.21. *For every choice of triangulation, the expression*

$$H_E^{T_\varepsilon}(N) = \frac{2}{9\beta} \sum_{\Delta \in T_\varepsilon} N(v(\Delta)) \sum_{IJK} \varepsilon^{IJK} \operatorname{tr}\left(A(\alpha_{IJ}(\Delta)) \left(A(p_K(\Delta))\right)^{-1} \{A(p_K(\Delta)), V(v(\Delta))\}\right) \tag{3.39}$$

tends to $H_E(N)$ *as* $\varepsilon \to 0$.

PROOF. According to the previous lemma to lowest order in ε we have

$$\begin{aligned}
&\frac{1}{\varepsilon^3} N(v(\Delta)) \sum_{IJK} \varepsilon^{IJK} \operatorname{tr}\left(A(\alpha_{IJ}(\Delta)) \left(A(p_K(\Delta))\right)^{-1} \{A(p_K(\Delta)), V(v(\Delta))\}\right) \\
=&\frac{1}{\varepsilon^3} N(v(\Delta)) \sum_{IJK} \varepsilon^{IJK} \operatorname{tr}\Big(\left(\mathbb{1} - \varepsilon^2 F_{ab}(v(\Delta)) u_I^a u_J^b\right) \left(\mathbb{1} + \varepsilon A_d(v(\Delta)) u_K^d\right) \\
&\qquad\qquad \times \{\mathbb{1} - \varepsilon A_c(v(\Delta)) u_K^c, V(v(\Delta))\} \Big) \\
=&\frac{1}{\varepsilon^3} N(v(\Delta)) \sum_{IJK} \varepsilon^{IJK} \Big(- \operatorname{tr}\left(\varepsilon \{A_c(v(\Delta)) u_K^c, V(v(\Delta))\}\right) \\
&\qquad\qquad - \operatorname{tr}\left(\varepsilon^2 A_d(v(\Delta)) u_K^d \{A_c(v(\Delta)) u_K^c, V(v(\Delta))\}\right) \\
&\qquad\qquad + \operatorname{tr}\left(\varepsilon^3 F_{ab}(v(\Delta)) u_I^a u_J^b \{A_c(v(\Delta)) u_K^c, V(v(\Delta))\}\right) \Big) \\
=&N(v(\Delta)) \sum_{IJK} \varepsilon^{IJK} u_I^a u_J^b u_K^c \operatorname{tr}\left(F_{ab}(v(\Delta)) \{A_c(v(\Delta)), V(v(\Delta))\}\right),
\end{aligned}$$

which is just the component of the three-form in the integrand of (3.33) (apart from a numerical factor). Note that we could drop the constant unit matrix from the Poisson bracket. The lowest-order term vanished because the trace of any single $\mathfrak{so}(3)$-matrix is zero, the next term canceled due to antisymmetry in I and J and the sum over those indices.

The statement then follows by writing (3.33) as a Riemann-sum. The factor $\frac{1}{\varepsilon^3}$ from the first line of the expression for the integrand cancels with the corresponding factor $\frac{\varepsilon^3}{6}$ from the volume of the tetrahedron. □

REMARK 3.22. As the proof shows, the factor $(A(p_K(\Delta)))^{-1}$ drops out in the limit, so one might think that this factor could be replaced by a unit matrix. However, it makes $H_E^{T_\varepsilon}(N)$ manifestly gauge-invariant since $\alpha_{IJ}(\Delta) \circ p_K^{-1}(\Delta) \circ p_K(\Delta)$ is a closed loop.

5.2. Quantization of the regulated classical constraint. As already mentioned, we quantize the Hamiltonian constraint by replacing all quantities in equation (3.39) by their operator analogs. The details are quite involved and we only briefly present the difficulties one runs into and their solutions. For a detailed discussion we refer to Thiemann's original article [37].

(1) Since we have already constructed the space $\mathcal{H}_{\text{diff}}$ of all solutions to the Gauß and diffeomorphism constraint, it would be desirable to define the Hamiltonian constraint as an operator on this space, i.e. as a linear function $\widehat{H}(N)\colon \mathcal{H}_{\text{diff}} \to \mathcal{H}_{\text{diff}}$.

The kernel of this operator would then naturally be equipped with an inner product (the restriction of the inner product on $\mathcal{H}_{\text{diff}}$) and we would immediately have a physical Hilbert space.

Unfortunately the Hamiltonian constraint maps diffeomorphism invariant states out of $\mathcal{H}_{\text{diff}}$, as is clear from the explicit form of the Dirac algebra (1.29b). To see this, suppose that $\Phi \in \mathcal{H}_{\text{diff}}$. Then (for the reason why the adjoint appears on the LHS see footnote 18)

$$\begin{aligned}\left((\widehat{H}^\dagger(N))'\Phi\right)\left(\widehat{V}(\phi)f\right) &= \Phi\left(\widehat{H}(N)\widehat{V}(\phi)f\right)\\ &= \Phi\left(\left[\widehat{H}(N),\widehat{V}(\phi)\right]f\right) + \Phi\left(\widehat{V}(\phi)\widehat{H}(N)f\right)\\ &= \Phi\left(-i\widehat{H}(\mathcal{L}_{M^c}N)f\right) + \Phi\left(\widehat{H}(N)f\right).\end{aligned}$$

Therefore $(\widehat{H}^\dagger(N))'$ is not in $\mathcal{H}_{\text{diff}}$, since the first term on the RHS does in general not vanish. Apart from complicating the construction of an inner product on the space of solutions, it is now impossible to investigate if the resulting theory is free of anomalies. For this purpose we would need to define the commutator (see equation (2.6)), which is impossible since we cannot apply the operator twice if the first application already leads out of its domain.

We circumvent this problem by implementing the Hamiltonian constraint as an operator on $\mathcal{H}_0$ and try to solve diffeomorphism and Hamiltonian constraint simultaneously. The solutions are then diffeomorphism invariant distributions Φ on $\mathcal{H}_0$ satisfying[17]

$$\Phi\left(\widehat{H}_E(N)f\right) = 0, \quad \forall f \in \mathcal{H}_0.$$

(2) Classically $H_E^{T_\varepsilon}(N)$ converges to $H_E(N)$ for every family of triangulations of Σ. We may therefore choose one such family, define the corresponding family $\widehat{H}_E^{T_\varepsilon}(N)$ and compute the limit $\varepsilon \to 0$ (in the sense explained below). The resulting operator therefore depends on the choice of triangulation, which means we are facing a serious quantization ambiguity. However, the allowed triangulations are restricted by physical considerations (such as e.g. that the resulting operator should not be trivial). Also diffeomorphism invariance removes a large part of the ambiguity, as explained briefly in [40].

Let us now describe one physically reasonable way of triangulating Σ. It turns out that the triangulation $T_\varepsilon(\gamma)$ has to be chosen dependent on the graph γ of the cylindrical function f_γ on which we want to define the action of the operator $\widehat{H}_E^{T_\varepsilon}(N)$.

(a) Every vertex and edge of γ shall also be contained in $T_\varepsilon(\gamma)$.

[17] Another solution is to combine all the Hamiltonian constraints for different lapse functions into a single master constraint, which actually does commute with the diffeomorphism constraint [39]. Rovelli on the other hand argues in favor of a Hamiltonian constraint , which *does* map diffeomorphism invariant states out of $\mathcal{H}_{\text{diff}}$ [32].

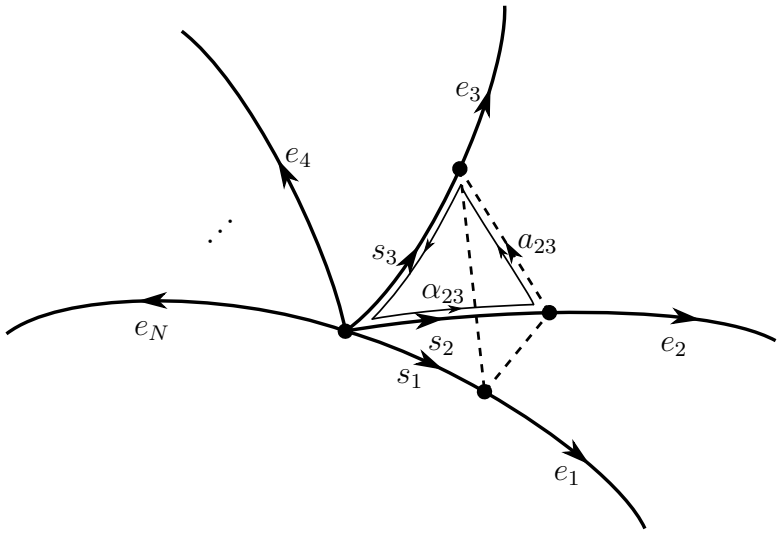

FIGURE 7. Illustration of a tetrahedron of the triangulation adapted to the graph of the SNWF.

(b) For each pair e, e' of edges of γ we choose an arc $a^{\varepsilon}_{\gamma(s),v,e,e'}$ going from an interior point of e to an interior point of e' and not intersecting γ in any other point. The segment of e from v to the intersection point with the arc will be denoted by $s^{\varepsilon}_{\gamma,v,e}$ and the loop consisting of the two segments and the arc by

$$\alpha^{\varepsilon}_{\gamma,v,e,e'} := \left(s^{\varepsilon}_{\gamma,v,e'}\right)^{-1} \circ a^{\varepsilon}_{\gamma,v,e,e'} \circ s^{\varepsilon}_{\gamma,v,e}. \tag{3.40}$$

We require that arcs corresponding to the same pair of edges but different ε are diffeomorphic and also that segments $s^{\varepsilon}_{\varphi(\gamma),\varphi(v),\varphi(e)}$ (and arcs) corresponding to a diffeomorphic graph are diffeomorphic to the segments $\varphi(s^{\varepsilon}_{\gamma,v,e})$ (and arcs) that were moved by the same diffeomorphism.

(c) For every vertex v we choose a neighborhood $U^{\varepsilon}_{\gamma,v}$ that contains the arcs (3.40). The choice shall be such that neighborhoods for different vertices do not intersect and such that in the limit $\varepsilon \to 0$ they contract to the vertex v monotonically (the neighborhood $U^{\varepsilon}_{\gamma,v}$ is contained in $U^{\varepsilon'}_{\gamma,v}$ if $\varepsilon < \varepsilon'$).
We require that $U^{\varepsilon}_{\gamma,v}$ is triangulated by a set of tetrahedra $\Delta(\gamma, v, e, e', e'')$. Each tetrahedron is bounded by the segments and arcs just mentioned and there is one tetrahedron for each ordered triple of edges incident at v. The triangulation away from the vertices of γ is left unspecified since the action of $\widehat{H}^{T_\varepsilon}_E$ does not depend on it.

This description of the triangulation is certainly somewhat sloppy, but figure 7 should suffice to get the idea.

The action of the Hamiltonian constraint on a cylindrical function f_γ is then easy to write down. We just have to replace every quantity in (3.39) by its operator analog an write a commutator instead of the Poisson bracket. The only difference is, that the triangulation is now the special one just explained. In other words, $e_I(\Delta)$ are the segments $s^\varepsilon_{\gamma(s),v,e}$ and the loops $\alpha_{IJ}(\Delta)$ are the ones in (3.40). We drop a numerical prefactor and have[18]

$$\begin{aligned}\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger f_\gamma = \sum_{v\in V(\gamma)} N(v) \sum_{\Delta\colon v(\Delta)=v} \sum_{IJK} \varepsilon^{IJK} \operatorname{tr}\Big(A(\alpha_{IJ}(\Delta))A(p_K(\Delta)) \\ \times \left[(A(p_K(\Delta)))^{-1}, \widehat{V}(U_\varepsilon(v))\right]\Big) f_\gamma.\end{aligned} \tag{3.41}$$

The reader might ask if this family of operators is consistently defined. In other words if a function is cylindrical over more than one graph, the action of the Hamiltonian constraint operator has to be independent of the graph. This problem can be circumvented by defining the action only on SNWF. Since those form a basis, the resulting family of operators is automatically consistent.

(3) We now have a family $\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger_{\varepsilon>0}$ and the question arises if we can take the limit $\varepsilon \to 0$. Since already the volume operator is unbounded, the norm of every operator $\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger$ is infinite. Therefore we can rule out convergence in the usual operator norm topology.

Naturally one would then ask if the family of operators converges in the weak topology, in other words if $\lim_{\varepsilon\to 0}\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger f$ exists for every cylindrical function f. However, it can be shown that the limit in this topology is zero.

Surprisingly our family converges to a non-trivial operator $\left(\widehat{H}_E(N)\right)^\dagger$ in the following sense. For every $\Phi \in \mathcal{H}_{\text{diff}}$ and every cylindrical function f,

$$\Phi\left(\left(\widehat{H}_E(N)\right)^\dagger f\right) = \lim_{\varepsilon\to 0} \Phi\left(\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger f\right).$$

Actually, this can be seen from the fact that we chose our triangulations such that in the neighborhood of a vertex they are diffeomorphic for different ε. Intuitively, as ε decreases the arc (3.40) moves closer to the vertex. But we can then perform a scaling diffeomorphism that moves the arc back to its original location. Therefore the number $\Phi\left(\left(\widehat{H}^\varepsilon_E(N)\right)^\dagger f\right)$ is independent of ε and convergence is trivial! We have

$$\left(\widehat{H}_E(N)\right)^\dagger = \left(\widehat{H}^{\varepsilon_0}_E(N)\right)^\dagger$$

for an arbitrary (but sufficiently small) value $\varepsilon_0 > 0$. Notice also how physics enters in the definition of the topology through the usage of diffeomorphism invariant functionals.

[18]Since classically the Hamiltonian constraint is real-valued, it does not make a difference if we quantize H or $\overline{H}$. For technical reasons one chooses $\overline{H}$, which results in $H^\dagger$ instead of H.

(4) For the sake of completeness, we mention that this definition of the Hamiltonian constraint operator is free of anomalies. We refer to [37] for the proof that $\Psi\left(\left[\widehat{H}_E^\dagger(N), \widehat{H}_E^\dagger(N')\right]\right) = 0$.

(5) Finally let us mention why the article [37] is called "Quantum Spin Dynamics". The notion stems from the intuitive picture of the theory: $\widehat{H}_E$ acts on SNWF T_s by changing the spin quantum numbers corresponding to the edges of the graph $\gamma(s)$.

Part 2

Loop Quantum Cosmology

Overview

Loop quantum cosmology is the study of highly symmetric cosmological models with methods of loop quantum gravity. In particular we want to emphasize that it is not the theory's cosmological sector. Let us mention briefly why the desirable goal of deriving cosmological solutions from full loop quantum gravity is not as easy a task as it may sound. In ordinary quantum mechanics symmetric models are treated by focusing on wave functions which are invariant under the action of a certain symmetry group. As spin networks represent the geometrical states of LQG, one might therefore try to look at spin networks which are invariant under the action of a certain symmetry group. However, this attempt is quickly seen to be unsuccessful. Any action of an element of the symmetry group (as for example a rotation) is in particular a diffeomorphism. Now in LQG we are dealing with a theory which is already invariant under *all* diffeomorphisms. Therefore this naive method does not restrict the allowed states at all.

Nevertheless, we are not stuck here if we recall from quantum mechanics that there are other ways of implementing symmetries. Having in mind e.g. rotationally invariant systems, one can choose coordinates adapted to the symmetry group and consider wave functions which depend only on selected coordinates (the radial coordinate r in the example of spherical symmetry). Mathematically this amounts to introducing δ-distributions for the remaining coordinates. Dealing with distributional states of course introduces new complications, but since the naive method from the previous paragraph fails in LQG this is our preferred choice.[19]

We therefore start with the same models as the popular mini-superspace quantizations of the Wheeler De-Witt theory. However, we will construct a representation different from the usual Schrödinger representation. Confidence that the resulting theory might nevertheless accurately describe the state of our universe near the classical "big-bang"-singularity is provided by the following observations: Firstly, the resulting states can be interpreted as distributional and in a certain sense symmetric states of the full theory [17], and secondly also the classical limit can be shown to be in good agreement with the Friedmann-Robertson-Walker models of our present universe [5]. Furthermore, even if

[19]Of course this method of employing distributional states is still not the desired cosmological sector of full LQG. A possible approach in this direction uses Rovelli's interpretation of the dynamics of LQG [32]. He starts with kinematical states $|\psi\rangle \in \mathcal{H}_{\text{diff}}$ and evolves them via a projector $\widehat{P}$ onto solutions of the Hamiltonian constraint. Physically one is then interested in transition amplitudes $\langle\psi_2|\,\widehat{P}\,|\psi_1\rangle$. To investigate cosmological questions one would then choose a coherent state ψ_1, peaked around a FLRW-like state.

these models do not accurately describe any physical situation, their study can provide insight for open problems of the full theory.

Cosmological Models

Let us start with a brief review of the cosmological models we are going to consider. In order to get a handle on the complex situation of the universe withing general relativity, symmetry principles have to be invoked. These are motivated by the Copernican principle, i.e. give rise to the requirements of homogeneity and isotropy. However, it is obvious that the observational data mankind has access to only describe a portion of our past light-cone. The assumptions of homogeneity and isotropy therefore contain a fair amount of philosophical input. But even though our universe is certainly not *exactly* homogeneous and isotropic, on sufficiently large scales our observational data are in agreement with those basic assumptions. Also the belief that our earth does not occupy a preferred position in space-time has been with us since the time of Copernicus and the assumption that there are no preferred directions in space also seems rather natural.

When quantizing symmetry-reduced models, it is often a good idea to embed them into a larger class. In this way more degrees of freedom get quantized and one can have increased confidence that the result might accurately describe the physical situation. Since a space-time which is isotropic at every point is necessarily also homogeneous (see e.g. [42]), a possible generalization that reflects some aspects of the universality is to study homogeneous but not isotropic models. Also from a physical perspective it is interesting to study cosmological models other than the homogeneous *and* isotropic ones. In this context we want to mention that the observation of accelerated expansion has in contrast to the interpretation as exotic matter prompted investigations that focus on averaged Einstein equations which differ from the simple FLRW-models [44].

In the context of the classical metric formulation, homogeneity means that space-time foliates into a preferred family of spacelike hypersurfaces Σ_t and that we are given an action of a Lie group S on space-time. The induced group-action on each Σ_t is required to be transitive and the metric has to be invariant under the action of each $s \in S$ (i.e. $s^* g_{ab} = g_{ab}$).

We will restrict our attention to models with a simply transitive group-action on Σ_t.[20] Homogeneous models with non-simply transitive action but with a simply-transitively acting subgroup can be treated as a semi-direct product (see below). It is interesting to note that the only homogeneous model which does not have a subgroup with simply transitive action is the Kantowski-Sachs model.

The assumption that the considered group-action is simply transitive implies that Σ_t is diffeomorphic to S. (Simply choose an arbitrary point $x_0 \in \Sigma_t$ and define a map $s \mapsto s(x_0)$). Therefore the task of writing down all homogeneous cosmological models

[20]The action is called transitive, iff for all $p, q \in \Sigma_t$ there is an element $s \in S$ such that $s(p) = q$. It is called simply transitive if this element is unique.

(except for the one mentioned above) reduces to finding all three-dimensional Lie groups, a task which was first accomplished by Bianchi in 1897. Homogeneous cosmological models with simply transitive action and without matter are therefore called Bianchi models.

The Einstein-equations for the metric then reduce to a set of ordinary differential equations. We refer to [42] for the details and proceed with our main interest, which is the study of Bianchi models in a connection formulation of general relativity. But before that we have to cover the general framework of symmetric principal fiber bundles.

CHAPTER 4

Symmetry Reduction

1. Symmetry Reduction for Theories of Connections

In the context of a gauge-theory formulation of general relativity, we have to consider symmetric principal fiber bundles [27, 17]. Contrary to the first part we therefore have to assume some knowledge of fiber bundles on the reader's part. The theory is presented in pedagogical form in [20].

For gravity we consider a principal fiber bundle $P(\Sigma, G, \pi)$ over a manifold Σ with structure group $G = SU(2)$. A bundle is called symmetric, if we are given an action $\Phi\colon S \to \mathrm{Aut}P$ of a Lie group S on P, such that every $\Phi(s)$ is a bundle automorphism. We remind the reader that a bundle morphism is map $f\colon P \to P$, which commutes with the right action of the bundle: $f(R_g p) = R_g f(p)$, $\forall g \in G$. In particular, bundle morphisms move whole fibers and therefore induce an action on the base manifold Σ. For simplicity we denote $\Phi(s)$ as well as the induced action on the base just by the group element s.

Since in cosmology we are only interested in homogeneous models, we will restrict our attention to actions which are fiber transitive (i.e. for every two fibers there is an element $s \in S$ which maps one fiber into the other; in other words, the induced action on the base Σ is transitive). Our presentation will closely follow [10], being more explicit at some points.

Let us start by classifying the set of all symmetric principal fiber bundles with transitive action of a given Lie group S. This will be interesting not only on its own, but also because we can then use the classifying elements to write down symmetric connections.

For that purpose we fix an arbitrary point $x_0 \in \Sigma$ and define the isotropy subgroup as the set of all elements of S which map the fiber over x_0 into itself, $F_{x_0} := \{s \in S \mid s(\pi^{-1}(\{x_0\})) \subseteq \pi^{-1}(\{x_0\})\}$. It is obvious that this is a subgroup of S and that for a different reference point $y_0 = t(x_0)$ we obtain the conjugated group $F_{y_0} = tF_{x_0}t^{-1}$. Being that conjugated groups are homeomorphic and diffeomorphic, we will write F instead of F_{x_0}, carefully keeping in mind that we chose a reference point and that our isotropy subgroup depends on this point!

The base space Σ of our principal fiber bundle is now isomorphic to S/F. To see this, note that $\varphi\colon S/F \to \Sigma,\ sF \mapsto s(x_0)$ is well-defined. Surjectivity follows from fiber-transitivity of the action while injectivity can be seen from the fact that $s(x_0) = t(x_0)$ implies $s^{-1}t \in F$ and therefore $sF = tF$.

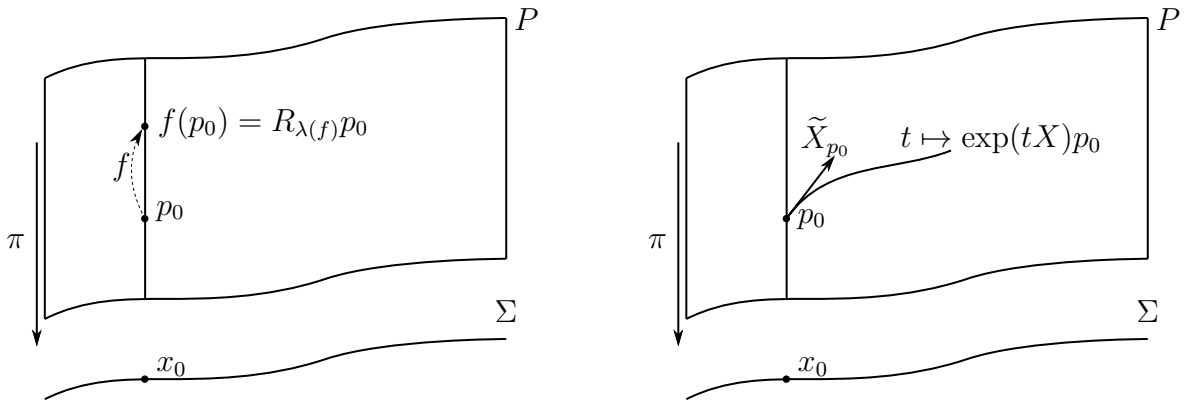

FIGURE 1. Definitions of λ_{p_0} and Λ_{p_0}.

The action of S also provides a homomorphism $\lambda\colon F \to G$ (see figure 1). We fix an arbitrary point p_0 in the fiber over x_0. Since for $f \in F$, $f(p_0)$ is again in the fiber over x_0, it can be expressed as a right translation of p_0 with a certain element $g \in G$. We define $\lambda(f)$ to be said element g. In other words, $f(p_0) =: R_{\lambda(f)}p_0$.[1] One quickly checks that λ is a group-homomorphism using commutativity of f_2 with right translations,

$$(f_2 f_1)(p_0) = f_2(R_{\lambda(f_1)}p_0) = R_{\lambda(f_1)}f_2(p_0) = R_{\lambda(f_1)}R_{\lambda(f_2)}p_0 = R_{\lambda(f_2)\lambda(f_1)}p_0,$$

from which we read off $\lambda(f_2 f_1) = \lambda(f_2)\lambda(f_1)$. It is also easy to see that a different reference point $R_g p_0$ leads to the conjugated homomorphism,

$$f(R_g p_0) = R_g f(p_0) = R_g R_{\lambda_{p_0}(f)}p_0 = R_g R_{\lambda_{p_0}(f)}R_{g^{-1}}R_g p_0 = R_{g^{-1}\lambda_{p_0}(f)g}R_g p_0,$$

giving $\lambda_{R_g p_0} = \mathrm{Ad}_{g^{-1}} \circ \lambda_{p_0}$.

The surprising result now is that the homomorphism λ suffices to reconstruct the principal fiber bundle. This is the content of the following

THEOREM 4.1. *For a given Lie group S all symmetric principal fiber bundles with fiber transitive action are characterized by a conjugacy class of homomorphisms $\lambda\colon F \to G$.*

PROOF. We just showed that every symmetric principal fiber bundle yields a conjugacy class of homomorphisms. Since the details of the reconstruction of $P(\Sigma, G, \pi)$ are not relevant for our further developments, we refer the reader to [18] for the general idea. The missing details should be easy to fill in. □

Since we now have control over all possible S-symmetric principal fiber bundles, we want to study invariant connections. Here a connection is called invariant if the differential $s'\colon TP \to TP$ maps every horizontal subspace onto a horizontal subspace for every $s \in S$. For the connection one-form this means $s^*\omega = \omega$.

[1] Obviously the map λ depends on the choice of the reference point p_0. We should therefore write λ_{p_0} to indicate this dependency. In order not to clutter our formulas with too much information, we choose to drop this index, carefully keeping in mind that λ depends on p_0. We will show in a moment, that a different reference point just yields a conjugated map, justifying this sloppiness in retrospect.

Given an invariant connection, we can use the connection one-form ω to construct a linear map[2]

$$\Lambda\colon \begin{cases} \mathfrak{s} \to \mathfrak{g} \\ X \mapsto \omega_{p_0}(\widetilde{X}) \end{cases} \tag{4.1}$$

with the fixed point p_0 from above. Here $\widetilde{X}$ is the vector field on P induced by $X \in \mathfrak{s}$. In other words, $\widetilde{X}_p$ is the tangent vector to the curve $t \mapsto \exp(tX)p$ at $t = 0$ (see figure 1). Note that for $X \in \mathfrak{f}$, i.e. the Lie algebra of the isotropy subgroup, we have that $\widetilde{X}_{p_0}$ is the tangent vector to $\exp(tX)p_0 = R_{\lambda(\exp tX)}p_0$, giving $\Lambda(X) = \lambda'(X)$ by the definition of the action of ω on a vertical vector. This part of Λ is therefore independent of the special connection under consideration. Λ satisfies the equation

$$\Lambda(\operatorname{Ad}_f X) = \operatorname{Ad}_{\lambda(f)} \Lambda(X), \quad \forall f \in F,\ X \in \mathfrak{s}. \tag{4.2}$$

Indeed, from $\exp(t \operatorname{Ad}_f X) = \operatorname{ad}_f \exp(tX)$ and the fact that $\exp(tX)$ acts as a bundle morphism we have

$$\begin{aligned} \exp(t \operatorname{Ad}_f X)p_0 &= (\operatorname{ad}_f \exp(tX))\, p_0 = f \exp(tX) f^{-1} p_0 \\ &= f \exp(tX) R_{\lambda(f^{-1})} p_0 = f \left(R_{\lambda(f^{-1})} \exp(tX) p_0 \right). \end{aligned}$$

This means $\widetilde{\operatorname{Ad}_f X}_{p_0} = f' R'_{\lambda(f^{-1})} \widetilde{X}_{p_0}$, which we can use to verify equation (4.2) (along with invariance of ω under the action of f),

$$\begin{aligned} \Lambda(\operatorname{Ad}_f X) &= \omega_{p_0}\left(\widetilde{\operatorname{Ad}_f X}\right) = \omega_{p_0}\left(f' R'_{\lambda(f^{-1})} \widetilde{X}\right) = \omega_{f^{-1}(p_0)}\left(R'_{\lambda(f^{-1})} \widetilde{X}\right) \\ &= \operatorname{Ad}_{\lambda(f)} \omega_{p_0}\left(\widetilde{X}\right) = \operatorname{Ad}_{\lambda(f)} \Lambda(X). \end{aligned}$$

The following theorem, which is due to Wang, states that the function Λ is sufficient to reconstruct the connection.

THEOREM 4.2 (Wang). *Let $P(\Sigma, G, \pi)$ be a symmetric principal fiber bundle with fiber transitive action of a connected Lie group S. Then the definition* (4.1) *establishes a 1:1 correspondence between the set of all symmetric connections on P and the set of all maps $\Lambda\colon F \to G$, which satisfy condition* (4.2) *and agree with λ' on vectors in $\mathfrak{f}$.*

PROOF. We already showed above that every symmetric connection yields a map Λ which satisfies the given conditions. We are therefore left with the task of reconstructing a connection out of a given map Λ, such that the map defined via (4.1) agrees with the original map.

For an arbitrary vector $Z \in T_pP$ in an arbitrary point $p \in P$ we have $p_0 = R_g s(p)$ for some $s \in S$, $g \in G$ since S acts fiber transitively. Therefore we get an element of $T_{p_0}P$

[2]Of course this map depends also on the reference point p_0 and we should therefore write Λ_{p_0} instead of Λ. Compare also footnote 1.

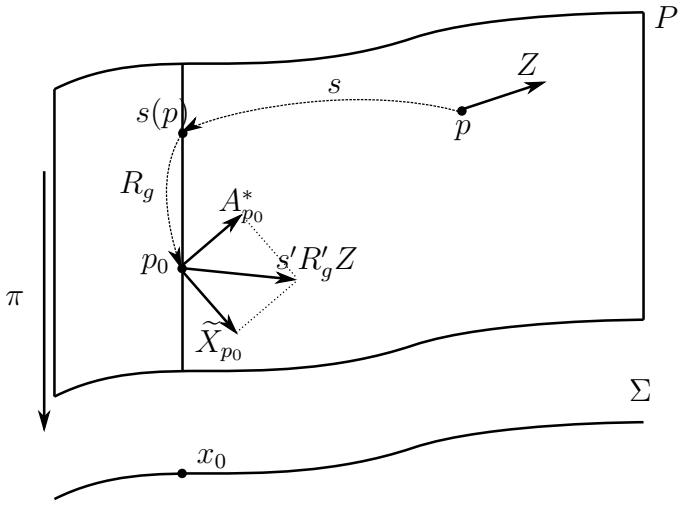

FIGURE 2. Illustration for the proof of theorem 4.2.

via $s'R'_gZ = \widetilde{X}_{p_0} + A^*_{p_0}$ with $X \in \mathfrak{s}$ and A^* is the fundamental vector field corresponding to $A \in \mathfrak{g}$ (see figure 2).[3]

From (4.1) and the known action of every connection on fundamental vector fields we know how ω has to act on this vector. Since we want to construct an invariant connection and since we know what a connection must do with a right-translated vector, we are forced to define

$$\omega(Z) := \mathrm{Ad}_g(\Lambda(X) + A).$$

It remains to show that this definition is independent of the chosen g and s, resp. X and A, and also that it satisfies the axioms of a connection and is invariant under the action of S. We will illustrate only the independence of the choice of X and A. The other calculations proceed along similar lines and can be found in [27].

Let therefore $\widetilde{X} + A^* = \widetilde{Y} + B^*$. Then $(B-A)^* = B^* - A^* = \widetilde{X} - \widetilde{Y} = \widetilde{X-Y}$. Acting on p_0 this equation yields $R_{\exp(t(B-A))}p_0 = \exp(t(X-Y))p_0 = R_{\lambda(\exp(t(X-Y)))}p_0$. Therefore $\lambda'(X-Y) = B - A$. On the other hand the second equation of this paragraph also shows that $\widetilde{X-Y}$ is vertical and therefore $X - Y \in \mathfrak{f}$. Thus, $\lambda'(X-Y) = \Lambda(X-Y) = \Lambda(X) - \Lambda(Y)$. Combining the last two equations yields $\Lambda(X) + A = \Lambda(Y) + B$. □

Since the map Λ has to agree with λ' on $\mathfrak{f}$, we pick a complementary space $\bar{\mathfrak{f}}$, so that $\mathfrak{g} = \mathfrak{f} \oplus \bar{\mathfrak{f}}$. It is convenient to assume that this can be done in such a way that this complement is invariant under the adjoint action of $\mathfrak{f}$, i.e. $\mathrm{Ad}_f(\bar{\mathfrak{f}}) \subseteq \bar{\mathfrak{f}}$.[4] In that case Λ is completely determined by its restriction to $\bar{\mathfrak{f}}$, which we call "Higgs"-field and denote by

[3]The fact that we can decompose every vector in $T_{p_0}P$ into a part which is induced by the action of S and a vertical vector again follows from fiber transitivity: The projection of $s'R'_gZ$ on Σ certainly can be integrated to an integral curve, $\pi'(s'R'_gZ) = \frac{d}{dt}\big|_{t=0} s_t x_0$. We define $\widetilde{X} := \frac{d}{dt}\big|_{t=0} s_t p_0$. This yields a vector with the same projection as $s'R'_gZ$. The difference $s'R'_gZ - \widetilde{X}$ is therefore vertical and corresponds to a fundamental vector field A^*.

[4] If this assumption is met we say that $\mathfrak{s}$ is reductive w.r.t. $\mathfrak{f}$.

$\Phi := \Lambda|_{\bar{\mathfrak{f}}}$. The condition (4.2) can be reformulated as

$$\Phi(\mathrm{Ad}_f X) = \mathrm{Ad}_{\lambda(f)}\, \Phi(X), \quad \forall X \in \bar{\mathfrak{f}}. \tag{4.3}$$

In LQG one works with gauge potentials. We will therefore have to pull back the connection one-form on P to a one-form on Σ by a canonical section. Of course we want to avoid any reference to the bundle itself and work with the classifying structure instead. We can do so by using sections $\iota\colon S/F \hookrightarrow S$ of the bundle $S(S/F, F)$ (i.e. embeddings $S/F \hookrightarrow S$ such that $\iota(sF) \in sF$). Any such section gives a section of the bundle P via $\sigma(y) := \iota(tF)(p_0)$, where $y = t(x_0)$. Now let $X \in T_y\Sigma$ with $y = s(x_0)$. Owing to transitivity of the action on Σ, X can be induced by a one-parameter group s_t, so that $X = \frac{d}{dt}\big|_{t=0} s_t s(x_0)$. To calculate the value $\omega(\sigma' X)$ we have to transport $\sigma' X = \frac{d}{dt}\big|_{t=0} \iota(s_t sF)p_0$ to our reference point p_0. Since this vector is tangent to P at the point $\iota(sF)p_0$, we have to act with the differential of $\iota(sF)^{-1}$,

$$\begin{aligned}(\iota(sF)^{-1})'\sigma' X &= \frac{d}{dt}\bigg|_{t=0} \big(\iota(sF)^{-1}\iota(s_t s)p_0\big) = \frac{d}{dt}\bigg|_{t=0} \big(\iota(sF)^{-1}\exp(t\iota' X)p_0\big) \\ &= \frac{d}{dt}\bigg|_{t=0} \big(\exp\big(t\Theta^S_{\mathrm{MC}}(\iota' X)\big)\, p_0\big) = \widetilde{\Theta^S_{\mathrm{MC}}(\iota' X)}_{p_0},\end{aligned}$$

where we implicitly used the isomorphism $\Sigma \to S/F$. Now we can easily read off the connection potential,

$$\sigma^*\omega = \Phi \circ \iota^*\Theta_{\mathrm{MC}}. \tag{4.4}$$

In classical symmetry reductions one usually chooses a fixed λ out of its conjugacy class. This leads to a breaking of the symmetry group from G to the centralizer of $\lambda(F)$ in G, which is defined by $Z_\lambda := Z_G(\lambda(F)) := \{g \in G \mid [g, \lambda(f)] = 0,\ \forall f \in F\}$. By that we mean the following. If Φ is a solution to the Higgs equation (4.3), then $\mathrm{Ad}_g \circ \Phi$ is another solution only if $g \in Z_\lambda$. Therefore, only gauge transformations with an element of Z_λ lead to a gauge potential which can again be written in the form (4.4). However, in passing to quantum theory we will want to have an $SU(2)$ connection.[5] One way to achieve this is by allowing arbitrary λ in the same equivalence class. It is easy to see that for a solution Φ to the Higgs equation with λ, $\mathrm{Ad}_{g^{-1}} \circ \Phi$ is a solution for $\mathrm{ad}_{g^{-1}} \circ \lambda$. Thus, going over from λ to $\mathrm{ad}_{g^{-1}} \circ \lambda$ induces a gauge transformation. It will however proof convenient later to work with a fixed λ. By also working with a fixed Higgs field compatible with λ we even fix the remaining Z_λ gauge freedom. The original $SU(2)$ gauge transformations can then be introduced "by hand" using the transformation $\Phi \mapsto \mathrm{Ad}_{g^{-1}} \circ \Phi$ with arbitrary $g \in SU(2)$. The transformed field will not satisfy the Higgs equation any more, but this does not pose a problem. In the next chapter we will introduce the basis τ_i in $\mathfrak{su}(2)$ and decompose $\Phi = \Phi^i \tau_i$. Gauge transformations then amount to a change of basis $\tau_i \mapsto g^{-1}\tau_i g =: \Lambda^j{}_i \tau_j$

[5] $SU(2)$ connections allow us to use the techniques developed for $SU(2)$ spin networks. Apart from the fact that this makes life easier because spin-network techniques for other gauge groups have not yet been fully developed, it is also appreciable because we want to test aspects of the full theory. A second point is of course that a selection of a fixed λ would lead to an unnatural breaking of the symmetry unless we take additional measures.

while keeping the components Φ^i fixed. Gauge transformations are therefore described by a matrix $\Lambda^i{}_j$ which satisfies the following equations.

LEMMA 4.3. *The matrix $\Lambda^i{}_j$ describing gauge transformations satisfies*

(1) $\Lambda^j{}_i\Lambda^j{}_k = \delta_{ik}$
(2) $\varepsilon^{ijk}\Lambda^l{}_i\Lambda^m{}_j\Lambda^n{}_k = \varepsilon^{lmn}$.

PROOF.

(1) By definition of $\Lambda^i{}_j$ we have $\operatorname{tr}(g^{-1}\tau_i g g^{-1}\tau_j g) = \Lambda^k{}_i\Lambda^l{}_j \operatorname{tr}(\tau_k\tau_l) = -\frac{1}{2}\Lambda^k{}_i\Lambda^k{}_j$. On the other hand, by cyclic invariance of the trace this expression is also given by $\operatorname{tr}(\tau_i\tau_j) = -\frac{1}{2}\delta_{ij}$.
(2) We consider $\operatorname{tr}(g^{-1}\tau_l g g^{-1}\tau_m g g^{-1}\tau_n g) = \Lambda^i{}_l\Lambda^j{}_m\Lambda^k{}_n \operatorname{tr}(\tau_i\tau_j\tau_k) = -\frac{1}{4}\Lambda^i{}_l\Lambda^j{}_m\Lambda^k{}_n\varepsilon_{ijk}$, which is also given by $\operatorname{tr}(\tau_l\tau_m\tau_n) = -\frac{1}{4}\varepsilon_{lmn}$.

□

For the convenience of the reader, let us summarize the findings of this section. In order to find all symmetric connections on symmetric principal fiber bundles with given symmetry group S and isotropy subgroup $F \leq S$, we start by finding all homomorphisms $\lambda\colon F \to G$. From each equivalence class (under conjugation) we take one element λ and find a solution to the Higgs equation (4.3). (We will always assume that $\mathfrak{s}$ is reductive w.r.t. $\mathfrak{f}$.) A local gauge potential is then given by (4.4), where $\iota\colon S/F \to S$ is such that $\iota(sF) \in sF$. Decomposing $\Phi = \Phi^i\tau_i$, gauge transformations are obtained by rotating $\tau_i \mapsto \Lambda^j{}_i\tau_j$, where $\Lambda^j{}_i$ satisfies the equations in lemma 4.3. This procedure will now be illustrated for cosmological models.

2. Bianchi Models in the Connection Framework

Specializing the framework of the preceding section to our cosmological models as in [10], the symmetry group S is now one of the Bianchi groups and the induced action on the base is simply transitive (in other words, $F = \{1\}$). Denoting the generators of $\mathfrak{s}$ by $\{T_I\}_{I=1,2,3}$, we have $[T_I, T_J] = c^I_{JK}T_K$ with the structure constants c^I_{JK} of $\mathfrak{s}$. For technical reasons, in a Hamiltonian treatment one can only treat Bianchi models of class A. By definition, their structure constants satisfy $c^J_{IJ} = 0$.

The assumption $F = \{1\}$ implies that there is only the trivial homomorphism $\lambda\colon F \to G,\ 1 \mapsto 1$. Also, because of $S/F = S$, there is only one section ι of the bundle $S(S/F, F)$, namely the identity on S. We decompose the Maurer Cartan form on S as $\Theta_{\mathrm{MC}} = \omega^I T_I$. The components of the Higgs field are defined by $\Phi(T_I) = \Phi^i_I\tau_i$. Since the Higgs condition is trivially satisfied, we arrive at our expression for the connection one-form of an invariant connection, $\Phi \circ \Theta_{\mathrm{MC}} = \Phi^i_I\tau_i\omega^I$. In the notation used in LQG, we have $A^i_a = \Phi^i_I\omega^I_a$.

Since the canonically conjugate momenta have to satisfy the Poisson brackets (2.23), we also have to demand that $\widetilde{E}^a_i$ is invariant. Unfortunately we do not have the physical

metric at our disposal to provide a density weight for $\widetilde{E}^a_i$. We will therefore use the invariant metric $\mathring{q} := \sum_I \omega^I_a \omega^I_b$ for that purpose. The left-invariant vector fields corresponding to T_I will be denoted by X_I and satisfy $\omega^I(X_J) = \delta^I_J$. The set $\{\omega_{\mathring{q}} X^a_I\}_{I=1,2,3}$ then provides a basis for the left-invariant density-weighted vector fields. Therefore we can expand $\widetilde{E}^a_i$ in the form $\widetilde{E}^a_i = \omega_{\mathring{q}} p^I_i X^a_I$.

In terms of our reduced variables Φ^i_I and p^J_j the symplectic form reads[6]

$$\frac{2}{\beta}\int_\Sigma \mathbb{d}A^i_a \wedge \mathbb{d}\widetilde{E}^a_i = \frac{2}{\beta}\int_\Sigma \omega_{\mathring{q}} d\Phi^i_I \wedge dp^J_i \omega^I(X_J) = \frac{2V_0}{\beta} d\Phi^i_I \wedge dp^I_i,$$

with $V_0 := \int_\Sigma \omega_{\mathring{q}}$. We choose to absorb V_0 into the Immirzi parameter and define $\beta' := \beta V_0^{-1}$. The symplectic form then corresponds to the Poisson brackets

$$\left\{\Phi^i_I, p^J_j\right\} = \frac{\beta'}{2}\delta^i_j \delta^J_I.$$

We have already mentioned that models with non-trivial isotropy subgroup $F \neq \{1\}$ can be treated as semi-direct products $S = N \rtimes_\rho F$ of a Bianchi group N with the isotropy subgroup F. Since the homomorphism λ and the Higgs equation depend on F and the representation ρ, we will first examine the Maurer Cartan form and the possible sections ι. For that purpose, we can keep the abstract notation. Recalling the composition for semi-direct products, $(n_1, f_1)(n_2, f_2) := (n_1\rho(f_1)n_2, f_1 f_2)$ we obtain the formula for inverse elements, $(n, f)^{-1} = (\rho(f^{-1})n^{-1}, f^{-1})$. We can use this to show that the Maurer Cartan form on S decomposes into the Maurer Cartan forms on N and F,

$$\begin{aligned}\Theta^S_{\mathrm{MC}}(n, f) &= (n, f)^{-1} d(n, f) = \left(\rho(f^{-1})n^{-1}, f^{-1}\right)(dn, df) \\ &= \left(\rho(f^{-1})n^{-1}\rho(f^{-1})dn, f^{-1}df\right) = \left(\rho(f^{-1})n^{-1}dn, f^{-1}df\right) \\ &= \left(\rho(f^{-1})\Theta^N_{\mathrm{MC}}(n), \Theta^F_{\mathrm{MC}}(f)\right).\end{aligned}$$

Due to $S/F = N$ all sections ι are given by maps $\iota\colon N \to S$, $n \mapsto (n, f(n))$. The simplest choice is of course $\iota\colon n \mapsto (n, 1)$ and we will do so. Our choice leads to the simple expression $\iota^*\Theta^S_{\mathrm{MC}} = \Theta^N_{\mathrm{MC}}$. Symmetric connections then are of the same form as before, $A^i_a = \Phi^i_I \omega^I_a$, where ω^I_a are still the components of the Maurer Cartan form on the Bianchi group N. The only difference is that now the Higgs equation is *not* trivially satisfied, which means we get a constraint on Φ.

Solving the Higgs constraint is different for every model. We will illustrate the procedure for isotropic models, where $F = SU(2)$ acts on N by rotations. It will be easier to solve the constraint (4.3) in its infinitesimal form. Since this equation has to hold for all $f \in F$, we can consider a curve $t \mapsto f_t$ in F with tangent vector τ_i at $t = 0$. Differentiation

[6] The careful reader might notice that the assumption of a compact spatial manifold Σ does not hold for the important case of the flat Bianchi I model, where Σ is isomorphic to $\mathbb{R}^3$. The definition of V_0 therefore fails, since an invariant metric cannot be integrated over a non-compact manifold. We can deal with non-compact Σ by introducing a fixed cell $\mathcal{V}$ which is assumed to be cubical with respect to the invariant metric $\mathring{q}_{ab}$. We will discuss how to deal with the unphysical volume V_0 of this arbitrary cell in chapter 6.

and evaluation at $t = 0$ then yields

$$\Phi\left(\mathfrak{ad}_{\tau_i}(X)\right) = \mathfrak{ad}_{\lambda'(\tau_i)}\,\Phi(X) = \left[\lambda'(\tau_i), \Phi(X)\right], \quad \forall X \in \bar{\mathfrak{f}},\ i = 1, 2, 3. \tag{4.5}$$

We have $\mathrm{ad}_{(1,f)}(n, 1) = (\rho(f)n, 1)$, implying that the adjoint action on the left hand side of this equation is the derivative of the representation ρ, which in our case is a rotation. For the basis elements T_I in $\bar{\mathfrak{f}} = \mathfrak{n}$, this term is therefore given by $\mathfrak{ad}_{\tau_i}(T_I) = \varepsilon_{iIK} T_K$. The previous equation simplifies to

$$\varepsilon_{iIK}\Phi^j_K = \Phi^k_I \left[\lambda'(\tau_i), \tau_k\right]_j, \quad \forall i, j, I = 1, 2, 3. \tag{4.6}$$

There are only two possible homomorphisms $\lambda\colon SU(2) \to SU(2)$. For the trivial homomorphism $\lambda_0(f) = 1$, the differential λ'_0 vanishes and so does the commutator on the right hand side of (4.6). The resulting Higgs condition $\varepsilon_{iIK}\Phi^j_K = 0$ only has the trivial solution $\Phi^i_I = 0$. Nevertheless, this case can serve to illustrate more elaborate concepts later on.

All other homomorphisms are given by $\lambda(g) = hgh^{-1}$ for an element $h \in G$. These are all conjugate and we make the simplest choice $\lambda_1 := \mathrm{id}$. With $\lambda'_1 = \mathrm{id}$ now the commutator becomes $\left[\tau_i, \tau_k\right]_j = \varepsilon_{ikj}$. The Higgs constraint reads $\varepsilon_{iIK}\Phi^j_K = \varepsilon_{ikj}\Phi^k_I$. The solutions are given by $\Phi^i_I = c\delta^i_I$ with an arbitrary parameter c. The momenta then have to be constant as well, $p^I_i = p\delta^I_i$ and the symplectic structure is $\frac{2}{\beta'} d\Phi^i_I \wedge dp^I_i = \frac{6}{\beta'} dc \wedge dp$.

CHAPTER 5

Original Approach to Loop Quantum Cosmology

Loop Quantum Cosmology was first studied by Bojowald in the Series of Papers [10, 11, 13, 14]. Motivated by success of general relativity in solving the Einstein equations in a symmetric context, we want to construct a representation of the symplectic space described in the previous chapter by mimicking as closely as possible the procedure of full LQG. Unfortunately the symmetry reduction does not only lead to simplifications. Some steps of the construction are also more complicated due to various reasons. The presentation is still along the lines of the just mentioned series of papers by Bojowald.

1. Kinematical Hilbert Spaces

Our configuration spaces consist of the linear maps Φ_I^i, showing that they have at most dimension 9 (for the anisotropic models). This is of course a great simplification as opposed to LQG, but it is also the first point where we cannot simply copy the procedure of the full theory. The Φ_I^i live just in the point x_0, the classifying structure of a symmetric principal fiber bundle has no direct reference to the spatial manifold Σ. Therefore, there is no direct way to construct holonomies $A(e)$, associated with an edge in Σ. Luckily, Thiemann has developed a formalism to deal with this issue in [36]. The idea is to exponentiate the three independent $\mathfrak{su}(2)$-valued components $\Phi_I = \Phi_I^i \tau_i$. The resulting $SU(2)$ objects $h_I := \exp(\Phi_I)$ are called point holonomies. Our kinematical Hilbert space is therefore $\mathcal{H}_{\text{kin}} = L^2\left(SU(2)^3, d\mu_H^3\right)$.

Unfortunately, now the conjugate momenta can not be represented as derivative operators on this space. To illustrate the problem, we use the following small

Lemma 5.1. *For a differentiable function $F\colon \mathbb{R} \to \mathfrak{g}$ the derivative of $\exp \circ F$ is given by*

$$\left.\frac{d}{ds}\right|_{s=s_0} \exp(F(s)) = \int_0^1 dt \exp((1-t)F(s_0))F'(s_0)\exp(tF(s_0)).$$

Proof. Defining the one-parameter family $U_s(t) := \exp\left(tF(s)\right)$, the LHS of our desired equality is given by $\delta U(t) := \left.\frac{\partial}{\partial s}\right|_{s=s_0} U_s(t)$ (for convenience we suppress the label s in $U_s(t)$). Differentiating by t we obtain

$$\dot{U}(t) := \frac{d}{dt}U(t) = F(s)\exp\left(tF(s)\right).$$

Variation of this equation yields a differential equation for $\delta U(t)$,

$$\frac{d}{dt}\left(\delta U(t)\right) = \delta\dot{U}(t) = \delta F U(t) + F\delta U(t).$$

The solution of the homogeneous equation is $\delta U(t)_{\text{hom}} = \exp(tF)\,C$. Variation of the constant C leads to the ansatz $\delta U(t) = \exp(tF)\,C(t)$ for the inhomogeneous equation. After inserting this ansatz we obtain

$$\dot{C}(t) = \delta F U(t),$$

and therefore

$$C(t) = \int_0^t dt' U^{-1}(t')\delta F U(t'),$$

where we set the integration constant to zero according to the boundary condition $\delta U(0) = 0$. Summing up we have

$$\delta U(t) = \int_0^t dt' U(t) U^{-1}(t')\delta F U(t').$$

Setting $t = 1$ and replacing t' by t we arrive at our desired result. □

To calculate the action of the momentum operators on our states, which are functions of the holonomies h_I, we specialize this lemma to the function $F\colon \Phi_I^i \mapsto \Phi_I^k \tau_k$. With $h_J = \exp(\Phi_J^j \tau_j)$ we obtain

$$\begin{aligned}\widehat{p}_i^I &= -\frac{i\beta'}{2}\frac{\partial (h_J)_B^A}{\partial \Phi_i^I}\frac{\partial}{\partial (h_J)_B^A}\\ &= -\frac{i\beta'}{2}\sum_{J=1}^{3}\delta_J^I\int_0^1 dt\left(\exp\left((1-t)\Phi_J^k\tau_k\right)\tau_i\exp\left(t\Phi_J^k\tau_k\right)\right)_B^A\frac{\partial}{\partial (h_J)_B^A}.\end{aligned}$$

Fortunately Thiemann has already found a way to deal with this continuous family of point holonomies. The solution consists in writing the point holonomy h_I as an actual holonomy along an edge e_I in an auxiliary manifold, along with a further regularization procedure. We will not go into the details but take the result from [36]. It basically amounts to inserting $\delta(t-0) + \delta(t-1)$ into the previous formula for $\widehat{p}_I^i$,

$$\widehat{p}_i^I = -\frac{i\beta'}{2}\sum_{J=1}^{3}\frac{1}{2}\left(\tau_i h_I + h_I\tau_i\right)_B^A\frac{\partial}{\partial (h_I)_B^A} = -\frac{i\beta'}{4}\left(R_{e_I,i} + L_{e_I,i}\right) = \frac{\beta'}{2}J_i^I. \tag{5.1}$$

The notation for the right and left invariant vector fields was introduced just above formula (3.15) and we abbreviated $-\frac{i}{2}\left(R_{e_I,i} + L_{e_I,i}\right) =: J_i^I$. Also notice the apparent similarity with that formula when e_I is an edge that starts and ends at x_0 and intersects a surface associated to "flux"-operator p_i^I at this point.

The regularization procedure is actually interesting in its own because writing point holonomies as actual holonomies provides us with good intuition on how to construct a loop representation. Let us therefore describe this part in more detail. We first have to introduce an auxiliary manifold for the edge e_I to lie in. It seems natural to reintroduce the base manifold S/F we already had before and take the integral curves to the left invariant vector fields X_I as regularizing edges. In other words, $\dot{e}_I(t) = X_I(e_I(t))$, where the vector fields X_I satisfy $\omega^J(X_I) = \delta_I^J$. The only problem is that we want our edges to be closed, starting and ending at x_0. Any other endpoint would introduce a preferred

point in S/F, unnaturally spoiling the symmetry. We therefore take a compactification $\overline{S/F}$ and assume that this can be done in such a way that the edges e_I are closed. Let us check that this holonomy indeed reproduces the point holonomy,

$$h(e_I) = \mathcal{P}\exp\left(\int_0^1 dt\dot{e}_i^a A_a^i \tau_i\right) = \mathcal{P}\exp\left(\int_0^1 dt\Phi^i{}_J \omega_a^J \dot{e}_I^a \tau_i\right) = \exp\left(\Phi_I\right). \tag{5.2}$$

We can therefore interpret our kinematical Hilbert space for the Bianchi models as consisting of spin network functions associated with spin networks consisting of the three edges e_I.

For the models with enhanced symmetry we also have to solve the Higgs equation. From the infinitesimal form (4.6) it is obvious that we get relations between the three holonomies e_I. The precise geometrical interpretation will be more easy to see if we exponentiate the original Higgs constraint (4.3),

$$\begin{aligned} h(f(e_I)) &:= \exp\left(\Phi(\mathrm{Ad}_f T_I)\right) = \exp\left(\mathrm{Ad}_{\lambda(f)}\Phi(T_I)\right) \\ &= \mathrm{Ad}_{\lambda(f)}\exp\left(\Phi(T_I)\right) = \mathrm{Ad}_{\lambda(f)}\Phi(T_I). \end{aligned} \tag{5.3}$$

The notation indicates that the left hand side can be interpreted as a holonomy along a rotated edge. However, the right hand side is not just the original holonomy — it also contains a "twisting" with $\mathrm{Ad}_{\lambda(f)}$.

For the isotropic model and the trivial homomorphism λ_0 the right hand side *is* just the unrotated holonomy $h(e_I)$. We already mentioned above that this is just a toy model, since there are no non-trivial solutions of the Higgs constraint.[1] Nevertheless, we can use it to gain some intuition. The configuration space of the anisotropic models is $SU(2)^3 = \{(h_1, h_2, h_3) \mid h_I \in SU(2)\}$. The constraint (5.3) demands that all three holonomies agree, reducing the configuration space to $\{(h, h, h) \mid h \in SU(2)\}$. This is a subgroup of $SU(2)^3$, which is isomorphic to $SU(2)$. The kinematical Hilbert space is therefore $\mathcal{H}_{\mathrm{kin}} = L^2(SU(2), d\mu_H)$. We can then use the Peter Weyl theorem to see that spin network functions with one edge form a basis.

Now let us examine the physically realistic isotropic model with the homomorphism $\lambda_1 = \mathrm{id}$, where the situation is complicated by the action of Ad_f on the right hand side of (5.3). The solution is easily found from the solution of the infinitesimal equation (4.6), which is equivalent to the exponentiated variant (the exponential map is locally one-one on compact, connected Lie groups). Exponentiation leads to the configuration space

$$\overline{\mathcal{A}}_{\mathrm{iso}} := \left\{\left(\exp(c\Lambda^i{}_1\tau_i), \exp(c\Lambda^i{}_2\tau_i), \exp(c\Lambda^i{}_3\tau_i)\right)\right\}, \tag{5.4}$$

where we have already undone the partial gauge-fixing. There are two points to mention about this configuration space. First, relaxing the partial gauge fixing introduces the parameters $\Lambda^j{}_i$. The treatment of the Gauß constraint therefore has to be adapted in order not to introduce additional degrees of freedom. We will come to this in the next

[1] With the exponentiated form of the Higgs constraint this can be seen quite easily. For a rotation by an angle of 180 degrees the left hand side is $h(e_I)^{-1}$, which has to equal $h(e_I)$. This is only possible for $h(e_I) = 1$, i.e. vanishing Higgs field.

section. The second point which complicates the construction is that this configuration space is no longer a subgroup of $SU(2)^3$. It is a union of equivalence classes under diagonal conjugation on $SU(2)$. To be precise, it can be written in the form $\cup_{c\in\mathbb{R}}\Theta_c$, where $\Theta_c := \{g(\exp(c\tau_1), \exp(c\tau_2), \exp(c\tau_3))g^{-1} \mid g \in SU(2)\}$. Therefore we can no longer use the Peter Weyl theorem to find a basis for the corresponding kinematical Hilbert space. We will construct a basis by hand later.

2. Gauß Constraint

We start by examining the Gauß constraint for the anisotropic models. In order to calculate (2.24) for our symmetric connections and densitized triads we can express the Ashtekar derivative in terms of another derivative and a difference tensor. Here it is natural to refer to the derivative $\mathring{D}$ defined by the Lie algebra $\mathfrak{su}(2)$. This derivative acts on tensors by acting with the vector field $\omega_a^I X_I$ on the coordinates of the tensor w.r.t. to the basis given by the one-forms ω_a^I and vector-fields X_I^a. In other words

$$\mathring{D}_a T^{b_1\ldots b_m}{}_{c_1\ldots c_n} := \left(X_I T^{I_1\ldots I_m}{}_{J_1\ldots J_n}\right)\omega_a^I X_{I_1}^{b_1}\ldots X_{I_m}^{b_m}\omega_{c_1}^{J_1}\ldots\omega_{c_n}^{J_n}.$$

Now we have to be careful since this derivative is not symmetric, i.e. the second derivatives do not commute on scalars. To see this we calculate

$$\begin{aligned}\left[\mathring{D}_a, \mathring{D}_b\right] f &= \left([X_I, X_J]\, f\right)\omega_a^I\omega_b^J = c_{IJ}^K\left(X_K f\right)\omega_a^I\omega_b^J \\ &= c_{IJ}^L\left(X_K f\right)X_L^c\omega_c^K\omega_a^I\omega_b^J = c_{IJ}^L X_L^c\omega_a^I\omega_b^J\mathring{D}_c f = T^c{}_{ab}\mathring{D}_c f,\end{aligned} \tag{5.5}$$

with the definition (N.4) of the torsion tensor. For the Gauß constraint we write

$$\mathcal{D}_a\widetilde{E}_i^a = \partial_a(\widetilde{E}_i^a) + \varepsilon_{ijk}A_a^j\widetilde{E}_k^a = p_i^I\partial_a\left(\omega_{\mathring{q}}X_I^a\right) + \varepsilon_{ijk}\Phi_J^j p_k^J\omega_{\mathring{q}} \tag{5.6}$$

with an arbitrary coordinate derivative ∂_a. In a second step we express this derivate in terms of $\mathring{D}_a$ and a difference tensor. We could of course skip this intermediate step and directly express $\mathcal{D}_a$ in terms of $\mathring{D}_a$, but in dealing with torsion it is easier to deal with derivatives that do not act on Lie-algebra indices.

For the *densitized* vector $\omega_{\mathring{q}}X_I^a$ our formula (N.2) is slightly modified. In general, for a densitized vector $\widetilde{v}^a$ we have

$$\begin{aligned}\nabla_a\widetilde{v}^b = \nabla_a v^b{}_{c_1\ldots c_n} &= \hat{\nabla}v^b{}_{c_1\ldots c_n} + C^b{}_{ca}v^c{}_{c_1\ldots c_n} - C^d{}_{c_1a}v^b{}_{dc_2\ldots c_n} - \cdots - C^d{}_{c_1a}v^b{}_{c_1\ldots c_{n-1}d} \\ &= \hat{\nabla}_a\widetilde{v}^b + C^b{}_{ca}\widetilde{v}^c - C^c{}_{ca}\widetilde{v}^b.\end{aligned}$$

Applying this general formula to our situation we have

$$\partial_a\left(\omega_{\mathring{q}}X_I^a\right) = \mathring{D}_a\left(\omega_{\mathring{q}}X_I^a\right) + \left(C^a{}_{ba} - C^a{}_{ab}\right)\omega_{\mathring{q}}X_I^b = -T^a{}_{ba}\,\omega_{\mathring{q}}X_I^b = -c_{IJ}^J\omega_{\mathring{q}},$$

where we made use of the fact that $\mathring{D}_a$ annihilates X_I^a and therefore also $\omega_{\mathring{q}}X_I^a$. The extra minus sign in the place where we used (N.5) stems from the fact that $C^a{}_{bc}$ is the difference of ∂_a w.r.t. $\mathring{D}_a$ and not vice versa. Inserting into (5.6) yields the Gauß

$$\widetilde{\mathcal{G}}_i = \omega_{\mathring{q}}\left(\varepsilon_{ijk}\Phi_I^j p_k^I - p_i^I c_{IJ}^J\right). \tag{5.7}$$

of the anisotropic cosmological models.

We will now show that the action of the corresponding operator on the kinematical Hilbert space is completely analogous to the action of the Gauß constraint in full LQG. For that purpose it is convenient make a parameter substitution. We agree that in the rest of this paragraph we will not sum over the indices $I, J, K, \ldots$ unless explicitly stated so. With the abbreviations

$$\begin{aligned} n_1^i(\alpha,\beta) &:= (\cos\beta, \cos\alpha\sin\beta, \sin\alpha\sin\beta) \\ n_2^i(\alpha,\beta) &:= (\sin\alpha\sin\beta, \cos\beta, \cos\alpha\sin\beta) \\ n_3^i(\alpha,\beta) &:= (\cos\alpha\sin\beta, \sin\alpha\sin\beta, \cos\beta) \end{aligned}$$

we substitute $\Phi_I^i =: r_I n_1^i(\alpha_I^1, \beta_I^1)$. Depending on the situation we may also substitute $\Phi_I^i =: r_I n_2^i(\alpha_I^2, \beta_I^2)$ or $\Phi_I^i =: r_I n_3^i(\alpha_I^3, \beta_I^3)$. We compute

$$\frac{\partial \Phi_I^k}{\partial \alpha_I^3} = r_I \left(-\sin\alpha_I^3 \sin\beta_I^3, \cos\alpha_I^3 \sin\beta_I^3, 0\right)^k = r_I \varepsilon_{3jk} n_3^j(\alpha_I^3, \beta_I^3).$$

Similar calculations for $i = 1, 2$ show that $\frac{\partial \Phi_I^k}{\partial \alpha_I^3} = \varepsilon_{ijk}\Phi_I^j$. By the chain rule we therefore have $\frac{\partial}{\partial \alpha_I^i} = \varepsilon_{ijk}\Phi_I^j \frac{\partial}{\partial \Phi_I^k}$, which is precisely the first term of the Gauß constraint. Since the second term in (5.7) vanishes for Bianchi class A models, the action of the dedensitized Gauß constraint on the kinematical Hilbert space is therefore given by (apart from numerical factors)

$$\widehat{\mathcal{G}}_i f = \varepsilon_{ijk}\Phi_I^j \frac{\partial f}{\partial \Phi_I^k} = \sum_{I=1}^{3} \frac{\partial f}{\partial \alpha_I^i}.$$

To calculate $\frac{\partial h_I}{\partial \alpha_I^i}$ we use a small trick, which we illustrate only for $i = 3$. We note that a gauge transformation with $\exp(\gamma\tau_3)$ transforms h_I into $\exp(\gamma\tau_3)\exp(r_I n_3^i(\alpha_I^3, \beta_I^3)\tau_i)\exp(-\gamma\tau_3) = \exp(r_I n_3^i(\alpha_i^3 + \gamma, \beta_I^3)\tau_i)$. In particular we can write the point holonomy as

$$h_I = \exp(\alpha_i^3\tau_3)\exp(r_I n_3^i(0, \beta_I^3)\tau_i)\exp(-\alpha_i^3\tau_3).$$

The derivative of the point holonomy is therefore $\frac{\partial h_I}{\partial \alpha_I^i} = \tau_i h_I - h_I \tau_i$. We thus arrive at

$$\widehat{\mathcal{G}}_i f = \sum_{I_1}^{3} (\tau_i h_I - h_I \tau_i)_B^A \frac{\partial f}{\partial (h_I)_B^A} = \sum_{I=1}^{3} \left(R_{e_I}^i - L_{e_I}^i\right) f.$$

This corresponds precisely to the action of the Gauß constraint in full LQG. The solutions can therefore be found in the same way as in LQG — a basis is given by gauge invariant spin network states on the three edges e_I with an intertwiner at the 6-valent (notice that every edge is closed and therefore incoming and outgoing) vertex x_0.

For models with enhanced symmetry the situation is not as straightforward. Classically (i.e. for fixed $\Lambda^j{}_i$) the gauge group is reduced to Z_λ. For the isotropic models $Z_\lambda = Z_{SU(2)}(SU(2)) = \{1\}$ is the center of $SU(2)$, which is trivial. The gauge group is completely broken and we expect a vanishing Gauß constraint. This can be confirmed by

inserting $\Phi^i_I = c\Lambda^i{}_I$ and $p^I_i = p_c\Lambda_i{}^I$ into (5.7),

$$\widetilde{G}_i = \omega_{\mathring{q}}\varepsilon_{ijk}cp_c\Lambda^j{}_I\Lambda_k{}^I = \omega_{\mathring{q}}\varepsilon_{ijk}cp_c\delta^j_k = 0.$$

However, this form of the constraint is irrelevant, since we already undid the gauge fixing. The gauge group is therefore $SU(2)$ and not $\{1\}$, and we have to enforce a full Gauß constraint for $SU(2)$ rotations also for the models with nontrivial isotropy subgroup.

These considerations are similar for models with isotropy subgroups other than $SU(2)$. Classically we get a Gauß constraint enforcing invariance under gauge transformations with elements of Z_λ. But after relaxing the gauge fixing we have to enforce a full $SU(2)$ Gauß constraint.

3. Diffeomorphism Constraint

We start in the same way as for the Gauß constraint, namely by deriving the reduced form of the constraint. Equation (2.25) shows that we need the expression for the curvature. We therefore insert the reduced form $A_a = \Phi_I\omega^I_a$ into (2.14). With the Maurer Cartan equations we obtain

$$\begin{aligned} F &= dA + \frac{1}{2}[A, A] = \Phi_I d\omega^I + \frac{1}{2}\omega^J \wedge \omega^K [\Phi_J, \Phi_K] \\ &= \frac{1}{2}\left(-\Phi^i_I c^I_{JK} + \varepsilon_{ijk}\Phi^j_J\Phi^k_K\right)\tau_i\omega^J \wedge \omega^K. \end{aligned}$$

Inserting into (2.25), smeared with a lapse function $\xi^a = \xi^I X^a_I$ (where ξ^I is constant due to homogeneity) we obtain

$$\begin{aligned} V(\xi) &= \int_\Sigma \widetilde{E}^b_i F^i_{ba}\xi^a = \int_\Sigma \omega_{\mathring{q}} p^L_i X^b_L \left(-\Phi^i_I c^I_{JK} + \varepsilon_{ijk}\Phi^j_J\Phi^k_K\right)\omega^J_{[b}\omega^K_{a]}\xi^M X^a_M \\ &= V_0 p^L_i \left(-\Phi^i_I c^I_{JK} + \varepsilon_{ijk}\Phi^j_J\Phi^k_K\right)\delta^{JK}_{ML}\xi^M = V_0\left(-\Phi^i_I c^I_{JK} p^K_i + \varepsilon_{ijk}\Phi^j_J\Phi^k_K p^K_i\right)\xi^J. \end{aligned}$$

Looking at (5.7) we see that the second term equals the Gauß constraint. Remembering our derivation of the diffeomorphism constraint in chapter 2, section 3, we can identify the first term as the actual cosmological diffeomorphism constraint

$$V(\xi) = -V_0 c^I_{JK}\Phi^i_I p^K_i \xi^J. \tag{5.8}$$

This expression looks surprisingly similar to the Gauß constraint (5.7), the only difference being that the structure constants ε_{ijk} of $SU(2)$ are now replaced by those of the symmetry group S. This raises the idea that the diffeomorphism constraint generates conjugation with elements of S in the base space Σ.

This is consistent with our geometric intuition for diffeomorphisms. In symmetry reduced models one uses special coordinates to exhibit the symmetry. Diffeomorphism are then limited to ones that preserve the special symmetric form of the coordinates. We saw in chapter 4 that the only freedom in the construction of symmetric coordinates lies in the choice of x_0. In our geometric picture the diffeomorphism constraint should therefore be related to this choice of the base point. To see the connection to (5.8) we note that

every point in Σ can uniquely be written in the form gx_0 with $g \in S$. The base manifold Σ is therefore identified with S via the map $gx_0 \mapsto g$. Performing a left translation with $s \in S$ yields a different base point sx_0. The point gx_0 is mapped to $sgx_0 = (sgs^{-1})sx_0$ and is therefore now identified with $\mathrm{ad}_s\, g$. We conclude that a different choice of reference point is equivalent to conjugation of the group S, which establishes to contact to (5.8).

For a more detailed discussion we have to specialize to a specific model. For the Bianchi I model the structure constants and therefore also the constraint (5.8) vanish. On the other hand also conjugation is trivial for commutative groups, so our two pictures are consistent. The diffeomorphism constraint for Bianchi I is trivially solved by using spin network functions on the three fixed edges as before.

Let us therefore discuss the more interesting Bianchi IX model. We first confirm for this particular model that the constraint (5.8) with an infinitesimal shift vector $\xi = \delta T_K$ really generates conjugation with $\exp(-\delta T_K)$. Our aim is therefore to find how $\mathrm{ad}_{\exp(-\delta T_K)}$ acts on a spin network function. For that purpose we note that the edges e_I of the auxiliary manifold get transformed $e_I^{(\delta,K)}(t) := \mathrm{ad}_{\exp(-\delta T_K)}\, e_I(t)$. They are tangent to $\dot{e}_I^{(\delta,K)}(t) = \mathrm{Ad}_{\exp(-\delta T_K)}\, \dot{e}_I(t)$. With Taylor's theorem we get

$$\mathrm{Ad}_{\exp(-\delta T_K)}\, T_I = T_I - \delta\, \mathfrak{ad}_{T_K}\, T_I + O(\delta^2) = T_I + \delta c_{IK}^L T_L + O(\delta^2),$$

which we can use to calculate the holonomy along the transformed edge. We remember that e_I is tangent to X_I, which corresponds to T_I and find

$$h_I^{(\delta,K)} = \mathcal{P} \exp \int_0^1 dt \Phi^i_J \omega^J \left(\mathrm{Ad}_{\exp(-\delta T_K)}\, X_I\right) \tau_i = \exp\left(\Phi^i_J \left(\delta_I^J + \delta c_{IK}^J + O(\delta^2)\right)\right) \tau_i.$$

Another Taylor expansion in δ shows that

$$\begin{aligned} h_I^{(\delta,K)} &:= \exp\left(\Phi^i_I \tau_i\right) + \delta \left.\frac{d}{d\delta}\right|_{\delta=0} \exp\left(\left(\varphi^i_I + \delta c_{IK}^L \Phi^i_L + O(\delta^2)\right) \tau_i\right) + O(\delta^2) \\ &= \exp\left(\Phi^i_I \tau_i\right) + \delta c_{MK}^L \Phi^j_L \frac{\partial}{\partial \Phi^j_M} \exp\left(\varphi^i_I \tau_i\right) + O(\delta^2) \end{aligned}$$

by the chain rule. One final Taylor expansion in δ show that for functions f of a holonomy we have

$$f(h_I^{(\delta,K)}) = f(h_I) + \delta c_{MK}^L \Phi^j_L \frac{\partial (h_I)^A_B}{\partial \Phi^j_M} f(h_I) + O(\delta^2).$$

We observe that the second term on the right hand side contains (up to constant factors) the quantization of the diffeomorphism constraint (5.8), which establishes the connections between diffeomorphisms and internal conjugations, resp. the choice of the reference point x_0.

The method to solve the diffeomorphism constraint is group averaging, just as in the full theory. This thereby also serves as illustration for the general procedure described in the first part (section 3.2). We start with a gauge invariant spin network function T_s of a spin network $s = \{(e_I)_{I=1,2,3}, (j_I)_{I=1,2,3}, i\}$ with arbitrary spins and an appropriate intertwiner v. As we have just seen, the cosmological diffeomorphism group acts

on the spin network by rotation. The orbit $\mathrm{Orb}(T_s)$ consists therefore of all $T_{s'}$ where $s' = \{(e'_I)_{I=1,2,3}, (j_I)_{I=1,2,3}, i\}$ is a spin network with rotated edges but same spins and intertwiner. The rigging map η then defines a basis for the diffeomorphism invariant states,

$$\eta(T_s)(T''_s) := \sum_{T_{s'} \in \mathrm{Orb}(T_s)} \langle T_{s'} T_{s''} \rangle \,.$$

Observe that because of orthogonality of the SNWF almost all terms in this sum vanish. The only exception is when the graph of s' equals that of s'' or at least contains the same edges in different order.

Similar to full LQG, where we defined knot classes, it is convenient to identify the linear functionals with equivalence classes $[\cdot]_{\mathrm{Orb}}$ of gauge invariant functions. We therefore have to find all possible orbits. The set of spin network functions on graphs with the three fixed edges we had earlier certainly encompasses all orbits, but this set is not independent: A spin network function which arises from a simple permutation of the three edges is in the same orbit as the original SNWF.[2] In order not to count a given SNWF multiple times we symmetrize this equivalence class and define $[T_s]_{\mathrm{Orb}} = \left[\frac{1}{\sqrt{6}} \sum_{\pi \in S_3} T_{\pi(s)}\right]_{\mathrm{Orb}}$ where S_3 is the permutation group and $\pi(s)$ is the spin network that arises from permuting the edges of s. Finally we observe that in the case that two spin labels are equal (e.g. $j_1 = j_2$), the spin network which has those two edges permuted gives rise to the same spin network function. We therefore do not need to symmetrize the corresponding edges and write $[T_s]_{\mathrm{Orb}} = \left[\frac{1}{\sqrt{3}} \left(T_s + T_{j_1 \leftrightarrow j_3} + T_{j_2 \leftrightarrow j_3}\right)\right]_{\mathrm{Orb}}$ and similarly for the case that the spin on two other edges coincides. For the case that all three spins are equal ($j_1 = j_2 = j_3$) no symmetrization is necessary.

In conclusion, the diffeomorphism invariant states are spanned by gauge invariant spin network functions on the three fixed edges. Solving the diffeomorphism constraint only adds the necessity of symmetrizing over edges which are labeled by different spins. We also observe the close relation to symmetry reduction for isotropic models. In both cases we are given an action of $SU(2)$ — but as we have seen symmetry reduction leads to a stronger reduction of degrees of freedom.

Speaking of isotropic models, we can insert $\Phi^i_I = c\Lambda^i{}_I$ and $p^K_i = p_c \Lambda_i{}^K$ into the diffeomorphism constraint (5.8) to see that it is proportional to c^I_{JI} and therefore vanishes identically for the Bianchi class A models we are considering. This of course was to be expected since we just saw that symmetry reduction is stronger than solving the diffeomorphism constraint. We can therefore simply work with spin networks of three fixed edges — no additional symmetrization is necessary.

[2]Actually, e.g. a permutation of the edges e_1 and e_2 can not be achieved by a rotation. But a rotation can transform (e_1, e_2, e_3) into $\left(e_2, e_1^{-1}, e_3\right)$. To our convenience this does not make a difference since for $SU(2)$ composing a representation with inversion leads to an equivalent representation. Therefore gauge invariant SNWF are not sensitive to inversion of an edge of the underlying graph.

4. Diffeomorphism invariant states for isotropic models

We have already seen that the configuration space for isotropic models is given by (5.4). In order to find all gauge invariant[3] (and therefore, according to the previous section, also diffeomorphism invariant) states, we could simply take the corresponding states of the anisotropic models and apply a pull-back with the canonical embedding $\overline{\mathcal{A}}_{\text{iso}} \to SU(2)^3$. In other words, we take gauge invariant functions on $SU(2)^3$ and restrict them to $\overline{\mathcal{A}}_{\text{iso}}$. However, since c carries the only gauge-invariant information, restrictions of gauge-invariant functions on $SU(2)^3$ can only depend on this one parameter. Our functions therefore depend on only one holonomy, the other two holonomies are redundant. Observe that this fits with our geometrical interpretation of the Higgs constraint, which roughly states that all three holonomies are equivalent. We see that the pull-backs of gauge invariant functions on $SU(2)^3$ are not independent. The purpose of this section is, apart from describing the measure μ_H on $\overline{\mathcal{A}}_{\text{iso}}$, to find an orthonormal basis of the space of gauge invariant functions in $L^2(\overline{\mathcal{A}}_{\text{iso}}, d\mu_H)$. We will try to give a good motivation and interpretation for the basis but the details of the proofs will be omitted and we refer to [11].

Our geometrical intuition would lead us to believe that all gauge invariant functions on $\overline{\mathcal{A}}_{\text{iso}}$ are given by gauge invariant spin network functions on a graph consisting of just one edge. For definiteness, let us choose the edge e_3 — then those functions will only depend on the holonomy $h_3 = \exp(c\Lambda_3)$, the other two edges will be invisible to this class of functions. However, this intuition is quickly seen to be incorrect. Gauge invariant spin network functions on a single edge are always invariant under the reflection $c \mapsto -c$. Geometrically we can easily understand this invariance when we remember that the adjoint action of $SU(2)$ on the group $SU(2)$ corresponds to a rotation in the Lie algebra.[4] Thus we have $g \exp(c\Lambda_3) g^{-1} = \exp\left(cR^I{}_3\Lambda_I\right)$ where $R^I{}_J$ is the rotation corresponding to $g \in SU(2)$. The statement that our considered space of functions is invariant under the reflection $c \mapsto -c$ then simply follows from the fact that there is a rotation which maps Λ_3 to $-\Lambda_3$ (e.g. a rotation of 180 degrees with axis Λ_1, which corresponds to an adjoint action with $g = \exp(\pi\Lambda_1)$). We can also verify this statement by a direct, algebraic calculation,

$$\exp(\pi\Lambda_1)\exp(c\Lambda_3)\exp(-\pi\Lambda_1) = -4\Lambda_1\exp(c\Lambda_3)\Lambda_1 = \exp(-4c\Lambda_1\Lambda_3\Lambda_1) = \exp(-c\Lambda_3)\,.$$

This establishes the statement that all gauge invariant functions on a single edge are invariant under $c \mapsto -c$. On the other hand, pull-backs of gauge invariant functions on $SU(2)^3$ can yield functions which are not invariant under this transformation. The reason is that for those functions the holonomies along the edges e_1 and e_2 are not "invisible" anymore. We would therefore need a gauge transformation with $h_I \mapsto h_I^{-1}$ for all $I =$

[3]In this context this means ad−invariant, i.e. invariant under $(h_1, h_2, h_3) \mapsto g(h_1, h_2, h_3)g^{-1}$ for all $g \in SU(2)$.

[4]This follows from $U\exp\left(v^i\tau_i\right)U^{-1} = \exp\left(v^iU\tau_iU^{-1}\right)$. Since for $U \in SU(2)$ we have that $U\tau_iU^{-1}$ is again traceless and anti-hermitian we can expand it in the basis τ_i for these matrices, $U\tau_iU^{-1} = L^j{}_i\tau_j$. That the linear transformation $L^j{}_i$ is actually a rotation follows from $\det\left(U\tau_iU^{-1}\right) = \det(\tau_i)$.

$1, 2, 3$. But since there is no rotation which maps Λ_I to $-\Lambda_I$ for $I = 1, 2, 3$ there can be functions which are not invariant under $c \mapsto -c$. Explicit examples are given by the basis functions ζ_j of theorem 5.2.

The missing class of functions can be motivated in the following way. A special set of gauge invariant spin network functions on the single edge e_3 is given by the traces $\operatorname{tr}(h_3^{m_3})$ where $h_3 = \exp(c\Lambda_3)$ and $m_3 \in \mathbb{Z}$.[5] Consider the application of the operator $2\Lambda^i{}_3 R_{e_3,i} = 2\left(\Lambda^i{}_3 \tau_i h_3\right)^A_B \frac{\partial}{\partial (h_3)^A_B}$ to these functions. We will see later that the volume operator contains this operator. Since the volume operator should preserve the kinematical Hilbert space, we are lead to assume that

$$2\Lambda^i{}_3 R_{e_3,i} \operatorname{tr}(h_3^{m_3}) = m_3 \operatorname{tr}\left(2\Lambda^i{}_3 \tau_i h_3^{m_3}\right) = m_3 \operatorname{tr}\left(\exp(\pi\Lambda_3) h_3^{m_3}\right) \tag{5.9}$$

is also contained in the kinematical Hilbert space. Observe that a gauge-transformation with $g = \exp(\pi\Lambda_1)$ transforms $\exp(\pi\Lambda_3)h_3$ into $-\exp(\pi\Lambda_3)h_3^{-1}$ (and h_3 simply into h_3^{-1} as already mentioned above). Since the trace is invariant under a gauge transformation we have

$$\operatorname{tr}\left(\exp(\pi\Lambda_3)h_3^{m_3}\right) = \operatorname{tr}\left(-\exp(\pi\Lambda_3)h_3^{-m_3}\right) = -\operatorname{tr}\left(\exp(\pi\Lambda_3)h_3^{-m_3}\right),$$

showing that the new functions pick up an additional sign under the mapping $c \mapsto -c$. They can therefore not be expanded in the original functions $\operatorname{tr}(h_3^{m_3})$ (invariance properties obviously carry over to linear combinations) and form a new class of gauge invariant functions on $\overline{\mathcal{A}}_{\text{iso}}$. We can also give a nice geometrical picture for our newly obtained class of functions. Having the same form as the original ones with only an additional factor $\exp(\pi\Lambda_3)$ inserted into the trace, we can interpret them as gauge-invariant spin network functions on the edge e_3 with an additional insertion in the vertex. Of course it is not at all clear that those two classes of functions already span the gauge invariant subspace of $L^2(\overline{\mathcal{A}}_{\text{iso}}, d\mu_H)$. Neither is it a proof that our new functions are at all elements of this space.

We have already mentioned that ad-invariant functions on $\overline{\mathcal{A}}_{\text{iso}}$ only depend on the parameter c and not on Λ_I. It follows that any such function can be written as $f(h_1, h_2, h_3) = F(h_3)$ with a function F on $SU(2)$.[6] The Ashtekar measure μ_0 of full LQG reduces to the Haar measure μ_H^3 on $SU(2)^3$. We refer to [11] for the statement that when further reduced to the subspace $\overline{\mathcal{A}}_{\text{iso}}$ of the isotropic models we obtain a measure which corresponds (in the sense of the Riesz representation theorem D.4) to the functional

$$F \mapsto \int_{SU(2)} F(h_3) d\mu_H(h_3)$$

[5]In fact it can be shown that these functions already form a basis of the gauge invariant spin network functions on single edge using the theorem of Peter and Weyl D.2 and Schur's lemma D.1.

[6]In the following we will always denote gauge invariant functions on $\overline{\mathcal{A}}_{\text{iso}}$ by the corresponding function F on $SU(2)$.

with the usual Haar measure on $SU(2)$. Integrating over the parameter Λ_3 yields the functional[7]

$$F \mapsto \frac{1}{2\pi}\int_0^{4\pi} F(c)\sin^2\frac{c}{2}dc.$$

Written as functions of the real parameter c the gauge invariant spin network functions on a single edge correspond to

$$\chi_j(c) := \frac{\sin(j+\frac{1}{2})c}{\sin\frac{c}{2}}, \quad j \in \frac{1}{2}\mathbb{N}_0.$$

For the spin network functions with insertion we remember that the characters can be written as $\chi_j(c) := \operatorname{tr}\pi^{(j)}(g) = g_{A_1}{}^{(A_1}\dots g_{A_{2j}}{}^{A_{2j})}$, where $SU(2)$ is parameterized in the form $g = \exp(c\Lambda_3)$. We obtain

$$\begin{aligned}\xi_j(c) &:= \Lambda^i{}_3 X_i^3 \chi_j(c) = 2j\big(\Lambda_3\exp(c\Lambda_3)\big)_{A_1}{}^{(A_1}\dots\big(\exp(c\Lambda_3)\big)_{A_{2j}}{}^{A_{2j})}\\ &= \frac{d}{dc}\chi_j\left(\exp(c\Lambda_3)\right) = j\frac{\cos\left(j+\frac{1}{2}\right)c}{\sin\frac{c}{2}} - \frac{1}{2}\frac{\sin jc}{\sin^2\frac{c}{2}}.\end{aligned} \tag{5.10}$$

Unfortunately those functions are not orthogonal to the functions χ_j. The correct functions which yield an orthonormal basis can be motivated by the following consideration. The operator $\Lambda^i{}_3 X_i^3$ is not symmetric. From the inner product

$$\begin{aligned}\langle f, g'\rangle &:= \int_0^{4\pi}\overline{f(c)}g'(c)\sin^2\frac{c}{2}dc = -\int_0^{4\pi}\overline{f'(c)}g(c)\sin^2\frac{c}{2}dc\\ &- \int_0^{4\pi}\overline{f(c)}g(c)\sin\frac{c}{2}\cos\frac{c}{2}dc = -\int_0^{4\pi}\left(\left(\frac{d}{dc}+\cot\frac{c}{2}\right)\overline{f(c)}\right)g(c)\sin^2\frac{c}{2}dc,\end{aligned}$$

we read off the adjoint operator

$$\left(\frac{d}{dc}\right)^* = -\frac{d}{dc} - \cot\frac{c}{2}.$$

In the next chapter we will build a self-adjoint volume operator by employing the anti self-adjoint combination

$$\frac{1}{2}\left(\frac{d}{dc} - \left(\frac{d}{dc}\right)^*\right) = \frac{d}{dc} + \frac{1}{2}\cot\frac{c}{2}. \tag{5.11}$$

It is therefore natural to look into the functions ζ_j defined by

$$\zeta_j(c) := \left(j+\frac{1}{2}\right)^{-1}\left(\frac{d}{dc}+\frac{1}{2}\cot\frac{c}{2}\right)\chi_j(c) = \frac{\cos\left(j+\frac{1}{2}\right)c}{\sin\frac{c}{2}}, j \in \frac{1}{2}\mathbb{N}_0 \tag{5.12}$$

and

$$\zeta_{-\frac{1}{2}} := \left(\sqrt{2}\sin\frac{c}{2}\right)^{-1}.$$

For the proof of the following theorem we again refer to [11]. We only remark that the difficult part is to show completeness; the proof of orthonormality reduces to standard integrals that also appear in Fourier analysis.

[7]To be precise, when we abbreviate functions $F(h_3)$ which do not depend on Λ_3 by $F(c)$ we should denote this new function of a real variable by a new letter. We ignore this subtlety.

THEOREM 5.2. *An orthonormal basis of the kinematical Hilbert space* $L^2(\overline{\mathcal{A}}_{iso}, d\mu_H)$ *is given by*

$$\left\{\chi_j, \zeta_k \mid j \in \tfrac{1}{2}\mathbb{N}_0, k \in \tfrac{1}{2}\mathbb{N}_0 \cup \left\{-\tfrac{1}{2}\right\}\right\}.$$

5. Volume Operator in isotropic LQC

A necessary step towards finding the dynamics of LQC is a detailed discussion of the volume operator. We will see in this section that in the simplified domain of isotropic LQC it is indeed possible to compute its full spectrum. By now this is an easy task, since we will see that the orthonormal basis of the last section indeed consists of eigenfunctions of the volume operator.

We start with the classical expression (3.32) for the volume in full LQG. For cosmological models we can insert $\widetilde{E}^a_i = \omega_{\mathring{q}} p^I_i X^a_i$ and obtain

$$V = \int_\Sigma \sqrt{\tfrac{1}{3!}\omega^3_{\mathring{q}} \left|\varepsilon^{ijk}\underset{\sim}{\delta}_{abc} p^I_i p^J_j p^K_k X^a_I X^b_J X^c_K\right|}.$$

We can simplify this expression with the help of the identity $\det(X^a_I) = \left(\det\left(\omega^I_a\right)\right)^{-1} = \omega^{-1}_{\mathring{q}}$, which follows from $\omega^I_a \omega^I_b = \mathring{q}_{ab}$ with the multiplication theorem for determinants and $\det A = \det A^T$. This identity in turn implies $X^a_I X^b_J X^c_K \underset{\sim}{\delta}_{abc} = \omega^{-1}_{\mathring{q}} \varepsilon_{IJK}$ (proof by contraction with ε^{IJK}, which is sufficient since both sides are antisymmetric in I, J, K, so there is only one independent component to check). Using this in our last expression for the volume we arrive at

$$V = \int_\Sigma \omega_{\mathring{q}} \sqrt{\tfrac{1}{3!}\left|\varepsilon^{ijk}\varepsilon_{IJK} p^I_i p^J_j p^K_k\right|} = V_0 \sqrt{\tfrac{1}{3!}\left|\varepsilon^{ijk}\varepsilon_{IJK} p^I_i p^J_j p^K_k\right|}. \tag{5.13}$$

For isotropic models $p^I_i = p\Lambda^I{}_i$ and $\varepsilon_{IJK}\Lambda^I{}_i\Lambda^J{}_j\Lambda^K{}_k = \varepsilon_{ijk}$ lead to the remarkably simple expression

$$V = V_0 \left|p\right|^{\frac{3}{2}}. \tag{5.14}$$

Let us pause here for a moment and discuss an important aspect of quantization. In full LQG there is no operator corresponding to the densitized triad $\widetilde{E}^a_i$. We cannot quantize the expression (3.32) directly but have to rewrite it in terms of fluxes. In cosmology we *do* have access to the triad representation (see (5.1)) and it will be convenient to use it. In fact we *have* to use it, to make our cosmological models at all solvable. On the other hand, this is also one of the points where LQC differs from the full theory and it might be a possible starting point for improving the theory and curing some of its flaws. Since we follow Bojowald in this chapter, we will adopt his method of quantizing the intermediate equation (5.13). Note that the alternative approach discussed in the next chapter takes the simpler expression (5.14) as a starting point.

Let us therefore insert our expression (5.1) for $\widehat{p}^I_i$ into (5.13) to obtain

$$\widehat{V} = V_0 \left(\frac{\beta' l_p^2}{2}\right)^{\frac{3}{2}} \sqrt{\tfrac{1}{3!}\left|\varepsilon^{ijk}\varepsilon_{IJK} J^I_i J^J_j J^K_k\right|} = V_0 \left(\frac{\beta' l_p^2}{2}\right)^{\frac{3}{2}} \sqrt{\left|\varepsilon^{ijk} J^1_i J^2_j J^3_k\right|}. \tag{5.15}$$

We mention that this operator equals precisely a vertex contribution of the volume operator of the full theory of reference [6], which is certainly an argument in favor of using the triad representation in LQC. Unfortunately, despite the tremendous simplification that already occurred, it is still no easy task to find the full spectrum of the volume operator of an arbitrary Bianchi model. We will therefore reduce further to isotropic models, where the Higgs constraint provides a relation between the vector fields J^I.

For models with nontrivial isotropy the Higgs constraint provides relations of the form $h_I = gh_Jg^{-1}$. To find the corresponding relation between the vector-fields J^I we note that the component of the right-invariant vector-field $R^i_{e_I}\tau_i$ with respect to the basis-vector v^i of an orthonormal basis in $\mathfrak{g}$ can be written as

$$\operatorname{tr}\big(\big(v^i h_I\big)^T \frac{\partial}{\partial h_I}\big).$$

Writing the relation for the holonomies in the form $(h_J)^A{}_D = (g^{-1})^A{}_B (h_I)^B{}_C g^C{}_D$ we have

$$\frac{\partial}{\partial (h_I)^B{}_C} = \frac{\partial}{\partial (h_J)^A{}_D}\frac{\partial (h_J)^A{}_D}{\partial (h_I)^B{}_C} = \frac{\partial}{\partial (h_J)^A{}_D}(g^{-1})^A{}_B g^C{}_D = \big((g^{-1})^T \frac{\partial}{\partial h_J} g^T\big)^B{}_C,$$

leading to the following equation for the right-invariant vector-field

$$R^i_{e_I} = \operatorname{tr}\big(\big(\tau^i h_I\big)^T \frac{\partial}{\partial h_I}\big) = \operatorname{tr}\big(\big(\tau^i g h_J g^{-1}\big)^T (g^{-1})^T \frac{\partial}{\partial h_J} g^T\big) = \operatorname{tr}\big(\big(g^{-1}\tau_i g h_J\big)^T \frac{\partial}{\partial h_J}\big),$$

by using cyclicity of the trace and $(AB)^T = B^T A^T$. This shows that $R^i_{e_I}$ is the component of $R^j_{e_J}\tau_j$ w.r.t. the basis vector $g^{-1}\tau_i g$ of the rotated basis. In other words we have $R^j_{e_J}\tau_j = R^i_{e_I} g^{-1}\tau_i g$, or equivalently $R^i_{e_I}\tau_i = R^j_{e_J} g\tau_j g^{-1}$. Since the same derivation applies for the left-invariant vector-fields, we arrive at the desired relation

$$J^I = gJ^J g^{-1}, \text{ whenever } h_I = gh_Jg^{-1}.$$

Using this general formula for our isotropic models, where the Higgs constraint yields (alternatively this can also be verified by a direct calculation)

$$\begin{aligned} h_1 &= \exp\left(\tfrac{\pi}{2}\Lambda_2\right) h_3 \exp\left(-\tfrac{\pi}{2}\Lambda_2\right) \\ h_2 &= \exp\left(-\tfrac{\pi}{2}\Lambda_1\right) h_3 \exp\left(\tfrac{\pi}{2}\Lambda_1\right) \end{aligned}$$

leads to the relations

$$\begin{aligned} J^1 &= \exp\left(\tfrac{\pi}{2}\Lambda_2\right) J^3 \exp\left(-\tfrac{\pi}{2}\Lambda_2\right) \\ J^2 &= \exp\left(-\tfrac{\pi}{2}\Lambda_1\right) J^3 \exp\left(\tfrac{\pi}{2}\Lambda_1\right). \end{aligned}$$

The essential ingredient of the volume operator (5.15) is $\varepsilon^{ijk} J^1_i J^2_j J^3_k$, which can also be written as $-4\operatorname{tr}(J^1 J^2 J^3)$ using $\operatorname{tr}(\tau_i\tau_j\tau_k) = -\frac{1}{4}\varepsilon_{ijk}$. It contains terms of the form $\exp\left(\frac{\pi}{2}\Lambda_I\right)\tau^i \exp\left(-\frac{\pi}{2}\Lambda_I\right)$, which we can simplify using $\exp\left(\frac{\pi}{2}\Lambda_I\right) = \frac{1}{\sqrt{2}} + \sqrt{2}\Lambda_I$ and

$\tau^i\tau^j\tau^k = \frac{1}{4}\left(\delta^{ik}\tau^j - \delta^{kj}\tau^i - \delta^{ij}\tau^k\right) - \frac{1}{8}\varepsilon^{ijk}$:

$$\begin{aligned}
\exp\left(\frac{\pi}{2}\Lambda_I\right)\tau^i\exp\left(-\frac{\pi}{2}\Lambda_I\right) &= \left(\frac{1}{\sqrt{2}} + \sqrt{2}\Lambda_I\right)\tau^i\left(\frac{1}{\sqrt{2}} - \sqrt{2}\Lambda_I\right) \\
&= \frac{1}{2}\tau^i + \Lambda_I\tau^i - \tau^i\Lambda_I - 2\Lambda_I\tau^i\Lambda_I \\
&= \frac{1}{2}\tau^i + \Lambda^j{}_I\varepsilon_j{}^i{}_k\tau^k - 2\Lambda^j{}_I\Lambda^k{}_I\frac{1}{4}\left(\delta^{jk}\tau^i - \delta^{ik}\tau^j - \delta^{ij}\tau^k\right) \\
&= \Lambda^j{}_I\varepsilon_j{}^i{}_k\tau^k + \Lambda^j{}_I\Lambda^i_I\tau_j.
\end{aligned}$$

Now we can simplify our main ingredient of the volume operator,

$$\begin{aligned}
-4\,\mathrm{tr}\left(J^1J^2J^3\right) &= -4J^3_iJ^3_jJ^3_k\,\mathrm{tr}\left(\exp\left(\frac{\pi}{2}\Lambda_2\right)\tau^i\exp\left(-\frac{\pi}{2}\Lambda_2\right)\tau^j\exp\left(-\frac{\pi}{2}\Lambda_1\right)\tau^k\exp\left(\frac{\pi}{2}\Lambda_1\right)\right) \\
&= -4J^3_iJ^3_jJ^3_k\,\mathrm{tr}\left(\left(\Lambda^l{}_2\varepsilon_l{}^i{}_m\tau^m + \Lambda^l{}_2\Lambda^i{}_2\tau_l\right)\tau^j\left(\Lambda^n{}_1\varepsilon_n{}^k{}_p\tau^p + \Lambda^n{}_1\Lambda^k{}_1\tau_n\right)\right) \\
&= J^3_iJ^3_jJ^3_k\,\mathrm{tr}\Big(\Lambda^l{}_2\Lambda^n{}_1\varepsilon_l{}^i{}_m\varepsilon_n{}^k{}_p\varepsilon^{mjp} - \Lambda^l{}_2\Lambda^n{}_1\Lambda^k{}_1\varepsilon_l{}^i{}_m\varepsilon^{mj}{}_n \\
&\qquad + \Lambda^l{}_2\Lambda^i{}_2\Lambda^n{}_1\varepsilon_n{}^k{}_p\varepsilon_l^{jp} - \Lambda^l{}_2\Lambda^i{}_2\Lambda^n{}_1\Lambda^k{}_1\varepsilon_l{}^j{}_n\Big) \\
&= J^3_iJ^3_jJ^3_k\,\mathrm{tr}\Big(\Lambda^l{}_2\Lambda^m{}_1\varepsilon_l{}^i{}_m\delta^{jk} - \Lambda^l{}_2\Lambda^j{}_1\varepsilon_l{}^{ik} + \Lambda^l{}_2\Lambda^l{}_1\Lambda^k{}_1\delta^{ij} \\
&\qquad - \Lambda^j{}_2\Lambda^i{}_1\Lambda^k{}_1 + \Lambda^l{}_2\Lambda^i{}_2\Lambda^l{}_1\delta^{jk} - \Lambda^k{}_2\Lambda^i{}_2\Lambda^j{}_1 - \varepsilon^{ijk}\Big) \\
&= J^3_iJ^3_jJ^3_k\left(\Lambda^i{}_3\delta^{jk}\right) = \Lambda^i_3J^3_i\left(J^3\right)^2.
\end{aligned}$$

Here, in going to the last line only the first term in the large bracket remained and was simplified using lemma 4.3 (2). All other terms vanished after using the first equation in said lemma, resp. after taking into account that $\Lambda^i{}_IJ^3_i$ vanishes for $I = 1, 2$ when acting on gauge-invariant states.[8] Putting this expression back into our previous result (5.15) we arrive at the volume operator

$$\widehat{V} = V_0\left(\frac{\beta' l_p^2}{2}\right)^{\frac{3}{2}}\sqrt{\left|\Lambda^i{}_3J^3_i\right|\left(J^3\right)^2} \tag{5.16}$$

of the isotropic models. However, we are still not satisfied with this operator. As already mentioned in the course of constructing an orthonormal basis for the gauge-invariant states, $\Lambda^i{}_3J^3_i$ is not self-adjoint. Since we clearly want a self-adjoint volume operator, we replace it with the self-adjoint combination $\frac{1}{2}\left(\Lambda^i{}_3J^3_i + (\Lambda^i{}_3J^3_i)^*\right)$. The corresponding derivative operator was already used before, it is $\frac{i}{3}$ times the operator in (5.11) and is to be regarded as quantization of the (real-valued) function $\Lambda^i{}_3\widetilde{E}^3_i$. Our final result for the

[8]This can be verified by considering the action on the basis states $\mathrm{tr}\left(h_3^{m_3}\right)$. Just as does $\Lambda^i{}_3J^3_i$, the operator $\Lambda^i{}_IJ^3_i$ leads to an insertion of $2\Lambda_I$ into the trace. But for $I = 1, 2$ the resulting trace vanishes, which is obvious since after expanding $h_3 = \cos\frac{c}{2} + 2\Lambda_3\sin\frac{c}{2}$ we are left with a linear combination of terms of the form $\mathrm{tr}\left(\Lambda_I(\Lambda_3)^n\right)$ which vanish thanks to $(\Lambda_3)^2$ being proportional to unity and lemma 4.3. Note that this argument is valid as well for the second class of basis states which already have an insertion of Λ_3.

self-adjoint volume operator of the isotropic models is

$$\widehat{V} = V_0 \left(\frac{\beta' l_p^2}{2}\right)^{\frac{3}{2}} \sqrt{\left|\widehat{\Lambda^i{}_3 \widetilde{E}_i^3}\right| \left(J^3\right)^2}. \tag{5.17}$$

For calculating the spectrum, we can now benefit from the work in section 4. From the definition (5.12) we immediately have

$$\widehat{\Lambda^i{}_3 \widetilde{E}_i^3} \chi_j = \frac{i}{3}\left(j + \frac{1}{2}\right)\zeta_j,$$

and simple differentiation shows that

$$\widehat{\Lambda^i{}_3 \widetilde{E}_i^3} \zeta_j = -\frac{i}{3}\left(j + \frac{1}{2}\right)\chi_j.$$

The operator

$$\left|\widehat{\Lambda^i{}_3 \widetilde{E}_i^3}\right| := \sqrt{\widehat{\Lambda^i{}_3 \widetilde{E}_i^3}^{\,2}}$$

therefore has eigenvalues $\frac{1}{3}\left(j + \frac{1}{2}\right)$ for $j > 0$, which correspond to the eigenfunctions χ_j and ζ_j. Furthermore, $\zeta_{-\frac{1}{2}}$ is eigenfunction with eigenvalue 0.

The other component of the volume operator, namely $\left(J^3\right)^2$ is identified with the derivative operator

$$\frac{1}{3^2}\left(\frac{d}{dc}\right)^* \frac{d}{dc} = \frac{1}{3^2}\left(-\frac{d^2}{dc^2} - \cot\frac{c}{2}\frac{d}{dc}\right).$$

It also has ζ_j and χ_j as eigenfunctions with eigenvalues $\frac{1}{3^2} j\left(j+1\right)$. For the functions χ_j this is best verified by differentiation using the trigonometric identity $-\cos\left(\left(j+\frac{1}{2}\right)c\right)\cos\frac{c}{2} = \sin\left(\left(j+\frac{1}{2}\right)c\right)\sin\frac{c}{2} - \cos jc$ (the function ξ_j was already computed in (5.10)),

$$\begin{aligned}\left(\frac{d}{dc}\right)^* \frac{d}{dc}\chi_j(c) &= \left(-\frac{d}{dc} - \cot\frac{c}{2}\right)\xi_j(c) = j\left(j+\frac{1}{2}\right)\frac{\sin\left(j+\frac{1}{2}\right)c}{\sin\frac{c}{2}} \\ &+ \frac{j}{2}\frac{\cos\left(j+\frac{1}{2}\right)c\cos\frac{c}{2}}{\sin^2\frac{c}{2}} + \frac{j}{2}\frac{\cos jc}{\sin^2\frac{c}{2}} - \frac{1}{2}\frac{\sin jc\cos\frac{c}{2}}{\sin^3\frac{c}{2}} \\ &- j\frac{\cos\left(j+\frac{1}{2}\right)c\cos\frac{c}{2}}{\sin^2\frac{c}{2}} + \frac{1}{2}\frac{\sin jc\cos\frac{c}{2}}{\sin^3\frac{c}{2}} = j\left(j+1\right)\chi_j(c).\end{aligned}$$

The corresponding identity for ζ_j can be verified in the same manner with little modifications. Alternatively, in order to save oneself this tedious work, one can use the result for χ_j and remember the definition (5.12). Once the identity

$$\left[\left(\frac{d}{dc}\right)^* \frac{d}{dc}, \frac{d}{dc}\right] = \left[\left(\frac{d}{dc}\right)^* \frac{d}{dc}, \left(\frac{d}{dc}\right)^*\right]$$

has been verified one then simply has

$$\begin{aligned}\left(\frac{d}{dc}\right)^* \frac{d}{dc}\zeta_j(c) &= \frac{2}{j+\frac{1}{2}}\left(\left(\frac{d}{dc}\right)^* \frac{d^2}{dc^2} - \left(\frac{d}{dc}\right)^* \frac{d}{dc}\left(\frac{d}{dc}\right)^*\right)\chi_j(c)\\ &= \frac{2}{j+\frac{1}{2}}\left(\frac{d}{dc}\left(\frac{d}{dc}\right)^* \frac{d}{dc} - \left(\frac{d}{dc}\right)^*\left(\frac{d}{dc}\right)^* \frac{d}{dc}\right)\chi_j(c)\\ &= \frac{2}{j+\frac{1}{2}}\left(\frac{d}{dc} - \left(\frac{d}{dc}\right)^*\right)\left(\frac{d}{dc}\right)^* \frac{d}{dc}\chi_j(c) = j\,(j+1)\,\zeta_j(c)\end{aligned}$$

Putting all pieces together, we see that the volume operator (5.17) of the isotropic models has eigenvalues

$$V_j = V_0\left(\frac{\beta' l_p^2}{2}\right)^{\frac{3}{2}}\sqrt{\frac{1}{27}j\left(j+\frac{1}{2}\right)(j+1)}, \quad j \in \frac{1}{2}\mathbb{N}_0. \tag{5.18}$$

For $j > 0$ they are degenerate with eigenfunctions ζ_j and χ_j, whereas the eigenvalue zero is threefold degenerate with eigenfunctions $\zeta_{-\frac{1}{2}}$, ζ_0 and χ_0.

6. Quantization of the Hamiltonian Constraint

We now take the last step towards the most interesting question of LQC — the discussion of its dynamics. We already mentioned that the open questions of full LQG are in this area and that one of the purposes of LQC is to gain insight into these problems. It is therefore mandatory to perform the quantization of the Hamiltonian constraint in close analogy to the construction presented in part 1.

However, this method makes use of certain features of the full theory which are not present in the symmetry-reduced context. Since in LQC the spatial manifold is reduced to a point, there is simply no room to lay the loop α_{IJ}, which was used to approximate the curvature components F^i_{ab}. Of course the solution will be to make use of the auxiliary manifold $\overline{S/F}$ that we already had to introduce before. Still, we have to be careful since we cannot just introduce arbitrary new loops but have to manage with the three edges e_I, $I = 1, 2, 3$. The kinematical Hilbert space of LQC simply does not account for other edges!

Our second challenge will be in a proper replacement for the continuum limit. Together with diffeomorphism invariance this limit allowed us to shrink the auxiliary loop to a point and thereby approximate F^i_{ab}. Recall from section 3 that there are cosmological models which don't even have a nontrivial diffeomorphism constraint! But let us not rush ahead and start our detailed presentation with the classical expression for the reduced Hamiltonian constraint.

The curvature components F^i_{IJ} were already computed in (3). Together with (2.22) we have

$$\widetilde{H}_E = \varepsilon_i{}^{jk} F^i_{IJ} p^I_j p^J_k \omega^2_{\mathring{q}} \omega_h^{-1} \quad \text{with} \quad F^i_{IJ} = -\Phi^i_K c^K_{IJ} + \varepsilon^i{}_{lm}\Phi^l_I \Phi^m_J. \tag{5.19}$$

Observe that the structure constants c^K_{IJ} of the symmetry group directly enter the Hamiltonian constraint. Thus, the dynamics will depend on the model under consideration.

For isotropic models the constraint simplifies considerably. For the flat model with $c_{IJ}^K = 0$ we have

$$\omega_{\mathring{q}}^{-2}\omega_h \widetilde{H}_E = \delta_{lm}^{jk}\delta_I^l\delta_J^m\delta_j^I\delta_k^J c^2p^2 = \delta_{lm}^{lm}\,(cp)^2 = 6\,(cp)^2\,, \tag{5.20}$$

whereas for Bianchi IX the structure constants are $c_{IJ}^K = \varepsilon_{IJ}{}^K$ and therefore

$$\begin{aligned}\omega_{\mathring{q}}^{-2}\omega_h \widetilde{H}_E &= -\varepsilon_i{}^{jk}\varepsilon_{IJ}{}^K\delta_K^i\delta_j^I\delta_k^J cp^2 + 6\,(cp)^2\\ &= -\varepsilon_i{}^{jk}\varepsilon_{jk}{}^i cp^2 + 6\,(cp)^2 = -6cp^2 + 6\,(cp)^2\,.\end{aligned} \tag{5.21}$$

The reformulation of the integrated constraint contained in lemma 3.19 can be transcribed to LQC with ease. We note that for homogeneous models the lapse function is constant. The cosmological version of (2.27) reads

$$\omega_h = \sqrt{\frac{1}{6}\left|\underset{\sim}{\delta}_{abc}\varepsilon^{ijk}\widetilde{E}_i^a\widetilde{E}_j^b\widetilde{E}_k^c\right|} = \omega_{\mathring{q}}\sqrt{\frac{1}{6}\left|\varepsilon^{ijk}\varepsilon_{IJK}p_i^I p_j^J p_k^K\right|}, \tag{5.22}$$

where we used $\underset{\sim}{\delta}_{abc}X_I^a X_J^b X_K^c = \varepsilon_{IJK}\omega_{\mathring{q}}^{-1}$ (write equation (A.13) as a determinant and use $\det A^{-1} = (\det A)^{-1}$).

Inserting into the constraint (5.19) we immediately see that the integrand of $H_E(N)$ only depends on Σ via $\omega_{\mathring{q}}$, which integrates to V_0. We therefore arrive at

$$H_E(N) = V_0 N\left(\frac{1}{6}\left|\varepsilon^{lmn}\varepsilon_{LMN}p_l^L p_m^M p_n^N\right|\right)^{-\frac{1}{2}}\varepsilon^{ijk}F_{IJk}p_i^I p_j^J.$$

In order to bring in the volume $V = V_0\sqrt{\frac{1}{6}\left|\varepsilon^{lmn}\varepsilon_{LMN}p_l^L p_m^M p_n^N\right|}$, we verify

$$\left\{\Phi_K^k, \varepsilon_{LMN}\varepsilon^{lmn}p_l^L p_m^M p_n^N\right\} = \frac{3}{2}\beta'\varepsilon^{ijk}\varepsilon_{MNK}p_i^M p_j^N$$

using our standard formula (C.7). The cosmological Hamiltonian constraint in terms of a Poisson bracket with the volume operator then reads

$$\begin{aligned}H_E(N) &= \frac{1}{3\beta'}V_0 N\left(\frac{1}{6}\left|\varepsilon^{lmn}\varepsilon_{LMN}p_l^L p_m^M p_n^N\right|\right)^{-\frac{1}{2}}\left\{\Phi_K^k, \varepsilon_{OPQ}\varepsilon^{opq}p_o^O p_p^P p_q^Q\right\}\varepsilon^{IJK}F_{IJk}\\ &= \frac{4}{\beta'}N\varepsilon^{IJK}\left\{\Phi_K^k, V\right\}F_{IJk}.\end{aligned} \tag{5.23}$$

6.1. Regularization. We now come to the discussion of the replacement for the regularization of the Hamiltonian constraint for cosmological models. In the full theory this was achieved by a triangulation of space. We approximated the curvature components by holonomies along closed loops and wrote the constraint as a Riemann sum over the tetrahedra of the triangulation. In order to perform the continuum limit diffeomorphism invariance was crucial. Unfortunately this procedure is not applicable in the cosmological context. The constraint is formulated at the point x_0 instead of on the whole manifold Σ, and we already discussed that the role played by diffeomorphisms differs greatly from full LQG.

The way to approximate the curvature components in LQC is to use the auxiliary manifolds $\overline{S/F}$. We have seen in (5.2) that they allow us to interpret the point holonomies as actual holonomies along the *closed* curves e_I. We can then lay a loop composed of those edges. Shrinking the loop to a point, we will then recover the curvature components. Unfortunately the required shrinking of the loop is troublesome here. Since the curves e_I have to be closed in order to identify a point holonomy with an actual holonomy along e_I, we cannot just take smaller pieces of e_I to create a smaller loop. The best we can do is to shrink the curve e_I by shrinking the whole manifold along with it. Since it is not clear how to shrink an abstract manifold without a natural metric at our disposal, we will try to gain some insight by considering the example of the flat Bianchi I model.

For the anisotropic Bianchi I model we have $S/F = \mathbb{R}^3$. The required compactification which renders the integral curves e_I to the invariant vector fields $X_I = \frac{\partial}{\partial x^I}$ closed is the three-torus. In other words, we identify the points $x^I = 0$ and $x^I = 1$. Writing the flow along the invariant vector fields in the form $t \mapsto \Phi_t(X_I)$ we have

$$\overline{S/F} = \left\{ \Phi_{t_3}(X_3)\Phi_{t_2}(X_2)\Phi_{t_1}(X_1) \mid t_I \in S^1 \right\}.$$

For this specific model it is obvious what we mean by a shrunk manifold. We simply identify the points $x^I = 0$ with $x^I = \varepsilon$ instead of with $x^I = 1$. This means that the parameter t_I now only runs from 0 to ε for a closed curve. Alternatively we can also keep $t_I \in [0, 1]$ and replace X_I by εX_I instead. We then have

$$\overline{S/F}_\varepsilon = \left\{ \Phi_{t_3}(\varepsilon X_3)\Phi_{t_2}(\varepsilon X_2)\Phi_{t_1}(\varepsilon X_1) \mid t_I \in S^1 \right\}. \tag{5.24}$$

Since this definition is applicable not only for Bianchi I, we will use it to define the shrinking manifolds for all cosmological models.[9]

From now on the quantization procedure is almost straightforward. We first have to extend the Hamiltonian constraint from the point x_0 to the auxiliary manifold $\overline{S/F}_\varepsilon$. The necessary integration yields just a factor of ε^3,

$$H_E^\varepsilon(N) := \int_{\overline{S/F}_\varepsilon} H_E(N) = \frac{4\varepsilon^3}{\beta'} N \varepsilon^{IJK} \left\{ \Phi_K^k, V \right\} F_{IJk}. \tag{5.25}$$

Looking at equation (5.19) we see that the curvature components depend explicitly on the symmetry group via the structure constants c_{IJ}^K. We therefore have to discuss our two models of interest separately.

[9]From a strictly topological standpoint, the three-torus $\overline{S/F}_\varepsilon$ is equivalent to $\overline{S/F}$. Thus we implicitly incorporate metric information to define the shrinking procedure (namely the metric of the underlying space $\mathbb{R}^3$. Therefore it is not at all clear if this method can be generalized to arbitrary symmetry groups. We have two points to say in our defense. Firstly, we are only interested in the Bianchi I and Bianchi IX models, and those are naturally embedded in an $\mathbb{R}^n$. Secondly, and more importantly, the regularizing manifolds are only needed in order to gain intuition and arrive at an expression of the Hamiltonian constraint in terms of point holonomies. This final expression will be independent of the shrinking manifolds and can be verified by a direct calculation without referring to our artificial construction. Nevertheless, the presented procedure creates a link to full LQG and provides a nice geometrical picture.

For the flat model, we can use a loop composed of the following four edges to approximate F_{IJ}: $c_1\colon t \mapsto \Phi_t(X_I)x_0$, $c_2\colon t \mapsto \Phi_t(X_J)c_1(1)$, $c_3\colon t \mapsto \Phi_t(-X_I)c_2(1)$, $c_4\colon t \mapsto \Phi_t(-X_J)c_3(1)$. From differential geometry we know that this parallelogram is closed, since for the flat model we have $[X_I, X_J] = 0$. Looking back at the calculation of (5.2) and taking into account $\Phi(\varepsilon X_I) = \varepsilon\Phi_I$ (linearity of the map Φ) we see that $h_{c_1}(A) = \exp(\varepsilon\Phi_I)$ and similarly for the other edges. Employing an expansion of the form $\exp(\varepsilon\Phi_I) = 1 + \varepsilon\Phi_I + \frac{1}{2}\varepsilon^2\Phi_I^2 + O(\varepsilon^3)$ yields

$$\begin{aligned}
&\exp(-\varepsilon\Phi_J)\exp(-\varepsilon\Phi_I)\exp(\varepsilon\Phi_J)\exp(\varepsilon\Phi_I)\\
&\quad= \left(1 - \varepsilon\Phi_J + \frac{1}{2}\varepsilon^2\Phi_J^2\right)\left(1 - \varepsilon\Phi_I + \frac{1}{2}\varepsilon^2\Phi_I^2\right)\left(1 + \varepsilon\Phi_J + \frac{1}{2}\varepsilon^2\Phi_J^2\right)\\
&\qquad\qquad\times\left(1 + \varepsilon\Phi_I + \frac{1}{2}\varepsilon^2\Phi_I^2\right) + O(\varepsilon^3)\\
&\quad= 1 - \varepsilon^2(\Phi_I\Phi_J - \Phi_J\Phi_I) + O(\varepsilon^3) = 1 - \varepsilon^2\varepsilon_{ijk}\Phi_I^i\Phi_J^j\tau^k + O(\varepsilon^3)\\
&\quad= 1 - \varepsilon^2 F_{IJ}^k\tau_k + O(\varepsilon^3)
\end{aligned} \tag{5.26}$$

for the holonomy around our closed loop.

For the Bianchi IX model the invariant vector fields do not commute. Hence the parallelogram will not be closed unless we also follow the curve $c_5\colon t \mapsto \Phi_t([X_I, X_J])c_4(1)$. We can use the result of the flat model to verify also for this model that the holonomy along the modified loop reproduces the curvature components,

$$\begin{aligned}
&\exp\left(\varepsilon^2 c_{IJ}^K\Phi_K\right)\exp(-\varepsilon\Phi_J)\exp(-\varepsilon\Phi_I)\exp(\varepsilon\Phi_J)\exp(\varepsilon\Phi_I)\\
&\quad= \left(1 + \varepsilon^2 c_{IJ}^K\Phi_K + O(\varepsilon^4)\right)\left(1 - \varepsilon^2\varepsilon_{ijk}\Phi_I^i\Phi_J^j\tau^k + O(\varepsilon^3)\right)\\
&\quad= 1 - \varepsilon^2\left(\varepsilon_{ijk}\Phi_I^i\Phi_J^j - c_{IJ}^K\Phi_K^k\right)\tau_k + O(\varepsilon^3) = 1 - \varepsilon^2 F_{IJ}^k\tau_k + O(\varepsilon^3).
\end{aligned}$$

Just as in full LQG we define

$$\begin{aligned}
H_E^\varepsilon(N) := \frac{8}{\beta'}N\sum_{IJK}\varepsilon^{IJK}\operatorname{tr}\Big(&\exp\left(\varepsilon^2 c_{IJ}^K\Phi_K\right)\exp(-\varepsilon\Phi_J)\exp(-\varepsilon\Phi_I)\\
&\exp(\varepsilon\Phi_J)\exp(\varepsilon\Phi_I)\exp(-\varepsilon\Phi_K)\left\{\exp(\varepsilon\Phi_K), V\right\}\Big).
\end{aligned} \tag{5.27}$$

It is easily verified that in the limit $\lim_{\varepsilon\to 0}\varepsilon^{-3}H_E^\varepsilon(N)$ we indeed reproduce the Hamiltonian constraint in the Form (5.23). Note that all leading terms in ε cancel for precisely the same reasons as in the full theory. Just as in the full theory we inserted a holonomy $\exp(-\varepsilon\Phi_K)$ right in front of the Poisson bracket. The limit would be correct also in the absence of this term, but we choose to add it for compliance with the full theory. We explicitly mention this fact here because in the original work [13] this holonomy was missing — however this was already corrected with the quantization of the inverse scale factor in [15] and the version with the added holonomy is also found in the subsequent paper [16].

Quantization of the expression (5.27) follows the usual procedure of replacing exponentials by multiplication operators with point holonomies and Poisson brackets by

commutators. We have

$$\widehat{H}_E(N) = -\frac{8i}{\beta'} N \sum_{IJK} \varepsilon^{IJK} \operatorname{tr}\left(h_{[I,J]} h_J^{-1} h_I^{-1} h_J h_I h_K^{-1} [h_K, \widehat{V}]\right), \tag{5.28}$$

where the quantization of the term involving the structure constants stems from

$$\exp\left(\sum_K c_{IJ}^K \Phi_K\right) = \prod_K \exp\left(c_{IJ}^K \Phi_K\right) \mapsto \prod_K (h_K)^{c_{IJ}^K} =: h_{[I,J]}.$$

Observe that when quantized as a multiplication operator the action of $\exp(\varepsilon\Phi_I)$ on diffeomorphism invariant states is actually independent of ε. This stems from our special construction of the shrinking procedure where we shrink the whole manifold and thereby also modify the holonomies and the states themselves. This feature fits perfectly with the situation of full LQG, where the action of the Hamiltonian constraint operator was also independent of the regulator when acting on diffeomorphism invariant states. It is therefore irrelevant if we formulate the action of the Hamiltonian constraint on $\overline{S/F}_\varepsilon$, or if we set ε to one and formulate it on $\overline{S/F}$. We can even forget about the auxiliary manifold altogether and regard (5.28) as an operator acting on functions of point holonomies (but remember that the introduction of an auxiliary manifold was also necessary for interpreting the point holonomies as actual holonomies).

By now we only discussed the Euclidean part of the Hamiltonian constraint. The full constraint can be treated in the same way as in full LQG. Again, one exploits the fact the integrated extrinsic curvature can be expressed as a Poisson bracket of the volume operator with the Euclidean constraint and uses the quantization (5.28) of the latter. We refer the interested reader to [13] for the details. This is admissible since we will subsequently focus on the isotropic flat model — and in this particular situation the full Hamiltonian constraint is proportional to the Euclidean part. Indeed, for flat models we have $dA = 0$, which means that the curvature simplifies to $F = \frac{1}{2}[A, A]$. Since also the spin connection Γ_a^i vanishes we have $A_a^i = \beta K_a^i$, leading to $F_{ab}^i = \varepsilon^i{}_{jk} A_a^j A_b^k = \beta^2 \varepsilon^i{}_{jk} K_a^j K_b^k$. The non-Euclidean term of the Hamiltonian constraint (2.22) is therefore

$$\begin{aligned}(H - H_E)(N) &= -\left(1 + \beta^2\right) \int_\Sigma N \varepsilon^{mij} \widetilde{E}_i^a \widetilde{E}_j^b \varepsilon_{mkl} K_a^k K_b^l \omega_h^{-1} \\ &= -\frac{1+\beta^2}{\beta^2} \int_\Sigma N \varepsilon^{mij} \widetilde{E}_i^a \widetilde{E}_j^b F_{ab}^m \omega_h^{-1} = -\left(1 + \frac{1}{\beta^2}\right) H_E(N),\end{aligned}$$

and the full constraint is proportional to the Euclidean part,

$$H(N) = -\frac{1}{\beta^2} H_E(N).$$

In order to specialize the Hamiltonian constraint to the now-considered isotropic models we insert $h_I = \exp(c\Lambda_I) = \cos\frac{c}{2} + 2\sin\frac{c}{2}\Lambda_I$ into our previous expression (5.28). We

have the intermediate expression

$$\begin{aligned} h_J^{-1}h_I^{-1}h_Jh_I &= \cos^4\frac{c}{2} + 2\left(1 - 2\varepsilon_{IJ}{}^K\Lambda_K\right)\sin^2\frac{c}{2}\cos^2\frac{c}{2} \\ &\quad + 4\left(\Lambda_I - \Lambda_J\right)\sin^3\frac{c}{2}\cos\frac{c}{2} - \sin^4\frac{c}{2}, \end{aligned} \tag{5.29}$$

which is valid for $I \neq J$ and follows easily by making use of this fact and of $\Lambda_I\Lambda_J = \frac{1}{2}\varepsilon_{IJ}{}^K\Lambda_K - \frac{1}{4}\delta_{IJ}$ (which is valid also for $I = J$. The second ingredient we need is

$$\begin{aligned} h_K^{-1}\left[h_K, \widehat{V}\right] &= \widehat{V} - \cos\frac{c}{2}\widehat{V}\cos\frac{c}{2} - \sin\frac{c}{2}\widehat{V}\sin\frac{c}{2} \\ &\quad + 2\Lambda_K\left(\sin\frac{c}{2}\widehat{V}\cos\frac{c}{2} - \cos\frac{c}{2}\widehat{V}\sin\frac{c}{2}\right), \end{aligned} \tag{5.30}$$

which is calculated along the same lines. When we insert those two expressions into (5.28) we see that the trace operation cancels all terms proportional to a single Λ_I. Terms proportional to unity and not depending on any of the indices I, J, K are canceled by the antisymmetric sum. There is therefore only one term which survives and which can easily be simplified with the help of $\mathrm{tr}\left(\Lambda_I\Lambda_J\right) = -\frac{1}{2}\delta_{IJ}$. We obtain (from now on we set the constant lapse function equal to one)

$$\widehat{H}(1) = \frac{192iV_0}{\beta^3}\sin^2\frac{c}{2}\cos^2\frac{c}{2}\left(\sin\frac{c}{2}\widehat{V}\cos\frac{c}{2} - \cos\frac{c}{2}\widehat{V}\sin\frac{c}{2}\right). \tag{5.31}$$

In order to find the action of this operator on a basis, the one we found in theorem 5.2 is rather inconvenient. The action of the operators $\cos\frac{c}{2}$ and $\sin\frac{c}{2}$ on the functions χ_j and ζ_j is not straightforward and it is favorable to introduce a new basis consisting of

$$|n\rangle := \frac{\exp\left(\frac{1}{2}inc\right)}{\sqrt{2}\sin\frac{c}{2}}, \quad n \in \mathbb{Z}. \tag{5.32}$$

The fact that those complex-valued functions form an orthonormal basis follows easily from the corresponding statement for χ_j and ζ_j and the fact that $|0\rangle = \zeta_{-\frac{1}{2}}$ and

$$|n\rangle = \frac{1}{\sqrt{2}}\left(\zeta_{\frac{1}{2}(|n|-1)} + i\,\mathrm{sgn}(n)\chi_{\frac{1}{2}(|n|-1)}\right).$$

From this equation we immediately see that the states $|n\rangle$ are also eigenstates of the volume operator. The eigenvalues are found by simply setting $j = \frac{1}{2}\left(|n| - 1\right)$ in (5.18),

$$\widehat{V}\,|n\rangle = V_0\left(\frac{\beta' l_P^2}{12}\right)^{\frac{3}{2}}\sqrt{\left(|n| - 1\right)|n|\left(|n| + 1\right)}\,|n\rangle =: \bar{V}_n\,|n\rangle \tag{5.33}$$

The action of the trigonometric operators on the new states is easily found using the identities

$$\begin{aligned} \cos\frac{c}{2} &= \frac{1}{2}\left(\exp\left(i\frac{c}{2}\right) + \exp\left(-i\frac{c}{2}\right)\right) \\ \sin\frac{c}{2} &= \frac{1}{2i}\left(\exp\left(i\frac{c}{2}\right) - \exp\left(-i\frac{c}{2}\right)\right). \end{aligned} \tag{5.34}$$

It is

$$\begin{aligned} \cos\frac{c}{2}\,|n\rangle &= \frac{1}{2}\left(|n+1\rangle + |n-1\rangle\right), \\ \sin\frac{c}{2}\,|n\rangle &= \frac{1}{2i}\left(|n+1\rangle - |n-1\rangle\right). \end{aligned} \tag{5.35}$$

With those equations it is now a cumbersome but easy exercise to compute the action of the Hamiltonian constraint operator (5.31) on the new states,

$$\widehat{H}(1)\,|n\rangle = \frac{6V_0}{\beta^3}\left(\bar{V}_{n+1} - \bar{V}_{n-1}\right)\left(|n+4\rangle - 2\,|n\rangle + |n-4\rangle\right). \tag{5.36}$$

This is a striking result since it shows that spatial discreteness (spectrum of the volume operator) in turn gives rise to temporal discreteness (spectrum of the Hamiltonian constraint).

Finally, let us comment on the question of anomalies in isotropic LQC. The fact that due to homogeneity we only use constant lapse functions means that Hamiltonian constraints commute, which is consistent with (1.29c). The remaining brackets of the hypersurface deformation algebra are fulfilled trivially since the diffeomorphism constraint of isotropic models vanishes.

CHAPTER 6

Simplified Kinematics

In [5] Bojowald, Ashtekar and Lewandowski realized that the kinematical structure of LQC can be simplified considerably. They present an alternative approach to the construction of the kinematical Hilbert space. Although it has several drawbacks (we will comment on them in the next chapter), the appeal of this construction lies in its simplicity and easier accessibility for beginners. This method is therefore also the preferred choice for non-specialists since it can also be presented without prior knowledge of Loop Quantum *Gravity* and can therefore serve as an introduction to the subject. Another advantage of the new approach is that the construction of the topology on the space of generalized connections (cf. section 2 in chapter 3), which is a very strong result of the full theory, can be carried over to cosmological models. To my knowledge this is not possible within the traditional approach of the previous chapter. Following the exposition in [5] we will present the new approach only in the regime of the flat isotropic model.

1. Improvements for the symmetry-reduced phase space

We start by briefly recapitulating and commenting on the symmetry reduction of chapter 4. Specialized to the now considered flat and isotropic models we have seen that connections and densitized triads can be written in the form

$$A_a^i = \breve{c}\omega_a^i, \quad \widetilde{E}_i^a = \omega_{\mathring{q}}\breve{p}X_i^a, \tag{6.1}$$

where we have written $\breve{c}, \breve{p}$ instead of c, p. The reason is that we will now modify those parameters in order to get rid of the unphysical volume V_0 of the fiducial cell $\mathcal{V}$ introduced in footnote 6 of chapter 4. Equation (6.1) expresses the connection one-form in terms of an invariant connection and the vector fields $\widetilde{E}_i^a$ in terms of invariant vector fields. Rescaling these invariant quantities by a constant yields again invariant quantities and we might just as well express our basic variables in terms of these rescaled quantities. In terms of the invariant metric defined by $\mathring{q}_{ab} := \omega_a^i\omega_b^i$ this rescaling amounts to $\mathring{q}_{ab} \mapsto k^2\mathring{q}_{ab}$ with an arbitrary constant $k \neq 0$. The one-forms then rescale according to $\omega_a^i \mapsto k\omega_a^i$. By duality, the invariant vector fields transform via $X_i^a \mapsto k^{-1}X_i^a$. Connection and densitized triads (6.1) can not be affected by the choice of an invariant metric in the class $\{k^2\mathring{q}_{ab} \mid k \in \mathbb{R} \setminus \{0\}\}$, since the dependence on such an unphysical structure would contradict general covariance. Thus , the parameter $\breve{c}$ transforms via $\breve{c} \mapsto k^{-1}\breve{c}$. The densitized triad also contains $\omega_{\mathring{q}}$, which yields a factor k^3 under rescaling. Therefore we have $\breve{p} \mapsto k^{-2}\breve{p}$.

Being coordinates of our physical phase-space, it is of course undesirable that $\breve{c}, \breve{p}$ depend on the unphysical choice of a fiducial metric. We can remedy this by defining

$$c := V_0^{\frac{1}{3}} \breve{c}, \quad p := V_0^{\frac{2}{3}} \breve{p}. \tag{6.2}$$

Those new quantities are not affected by rescalings since obviously $V_0 := \int_{\mathcal{V}} \omega_{\mathring{q}} \mapsto k^3 V_0$. The added benefit is that now the symplectic form does not depend on the parameter V_0. It reads

$$\frac{6}{\beta} dc \wedge dp$$

with an unmodified Immirzi parameter β. The connection and densitized triads of our isotropic models in terms of the new variables read

$$A_a^i = c V_0^{-\frac{1}{3}} \omega_a^i, \quad \widetilde{E}_i^a = \omega_{\mathring{q}} p V_0^{-\frac{2}{3}} X_i^a. \tag{6.3}$$

The modifications of this chapter only concern the quantization of the classical configuration space. Therefore all the previous results that were carried out on the classical level still apply. In particular the Gauß- and diffeomorphism constraints are solved and the remaining Hamiltonian constraint (with constant lapse function and integrated over the cell $\mathcal{V}$) reads

$$H(1) = -\frac{6}{\beta^2} c^2 \operatorname{sgn} p \sqrt{|p|} = -\frac{4}{\beta^3} \varepsilon^{ijk} \{c, V\} F_{abk} X_i^a X_j^b, \tag{6.4}$$

which can be transcribed from (5.20) and (5.23) using (5.22).

2. The cosmological equivalent of the holonomy-flux algebra

Contrary to the previous chapter we now do *not* work with point holonomies. The reduced connection (6.1) is now interpreted as an actual connection on a full $SU(2)$-bundle over the manifold Σ, instead of as a connection living on the reduced bundle over the point x_0. This opens the possibility to work with real holonomies from the start. We can therefore be closer to the full theory at this point of the construction — but this does not come without a price to pay as we will discuss in detail in the next chapter.

As in full LQG we choose our kinematical Hilbert space to consist of holonomies $h_e(A)$ along edges e in Σ. But owing to isotropy the requirement that the space of holonomies separates the points (cf. item 1 of section 1 in chapter 2) can now be fulfilled more easily: In order to recover isotropic connections it suffices to consider holonomies along edges which lie along straight lines. Here the term *straight* refers to the metric $\mathring{q}_{ab}$, in other words $\dot{e}^a \omega_a^i =: \Delta x^i$ is constant.[1]

[1]This follows easily from the geodesic equation: For a flat metric the compatible derivative is just the coordinate derivative. The geodesic equation therefore reads $\dot{e}^a \partial_a \dot{e}_b = 0$. Contracting with ω_b^i and taking into account $\partial_a \omega_b^i = 0$ then yields $0 = \dot{e}^a \partial_a \dot{e}_b \omega_b^i = \partial_t \left(\dot{e}^b(t) \omega_b^i(x(t)) \right)$, which means $\dot{e}^b \omega_b^i$ is constant.

We label our edges e by a parameter μ, which is defined in such a way that $\mu V_0^{\frac{1}{3}}$ is the oriented length of e with respect to $\mathring{q}_{ab}$. So, we have

$$\mu V_0^{\frac{1}{3}} := \int_0^1 \sqrt{\dot{e}^a(t)\dot{e}^b(t)\mathring{q}_{ab}(x(t))}dt = \sqrt{\Delta x^i \Delta x^i} =: \Delta x.$$

In order to compute the holonomy along an edge which is labeled by μ we start by observing that in

$$\int_e A_a = \int_0^1 \breve{c}\omega_a(x(t))\dot{e}^a(t)dt = \breve{c}\dot{e}^a\omega_a = \breve{c}\Delta x^i \tau_i,$$

the integrand is independent of the parameter t. The path ordering in the definition of the holonomy therefore drops out and the calculation reduces to taking exponentials of $\mathfrak{su}(2)$ matrices, a task which we are already familiar with at this point,

$$h_e(A) = \mathcal{P}\exp\int_e A = \cos\frac{\mu c}{2} + \frac{2\Delta x^i \tau_i}{\Delta x}\sin\frac{\mu c}{2}. \tag{6.5}$$

We saw in the proof of lemma 3.14 that the state space of LQG can be seen as the completion of the algebra of finite sums of products of matrix elements of holonomies. From (6.5) we see that in our cosmological setting this algebra is nothing else than the algebra of almost periodic functions, which consists of all functions of the form

$$g(c) := \sum_{\mu\in\mathbb{R}} \xi_\mu \exp\left(i\frac{\mu c}{2}\right), \tag{6.6}$$

where it is understood that only a finite number of the coefficients ξ_μ is nonzero. It is admissible to regard a finite number of edges, each labeled by a number $\mu \in \mathbb{R}$, as a graph and a function g of the form (6.6) as a cylindrical function over this graph. We will discuss this geometrical interpretation and its analogy with LQG in more detail below. For now, we just introduce the notation Cyl_S for the space of functions g, which we call *symmetric* cylindrical functions.[2]

Similar to the holonomies it is now also appropriate to admit only a subset of surfaces S and functions f^i in the definition of the fluxes $E_f(S)$. Due to homogeneity we only allow constant $\mathfrak{su}(2)$-valued functions f^i. The surfaces S will be restricted to squares whose sides are tangential to the vector fields X_i^a. With those assumptions we can calculate the fluxes,

$$E_f(S) := \int_S \mathrm{tr}\left(\widetilde{E}_i^a\right) f^i = pV_0^{-\frac{2}{3}} \int_S \mathrm{tr}\left(X_i^a \omega_{\mathring{q}}\right) f^i =: pV_0^{-\frac{2}{3}} A_{S,f},$$

where we defined $A_{S,f}$ proportional to the area of S, measured by $\mathring{q}_{ab}$. The momenta are therefore proportional to p and the analog of the holonomy-flux algebra (3.15) consists of

[2]The detailed discussion given below will show that it is more appropriate to call the functions T_μ SNWF. Hence the space Cyl_S consists of finite linear combinations of SNWF. Of course every such function is cylindrical over its graph, but they are not the only cylindrical functions in LQC. The term *space of* symmetric cylindrical functions for Cyl_S therefore does not seem fully appropriate and might lead to confusion. However, we follow the exposition in [5] here and keep their notation.

these momenta and the symmetric cylindrical functions. The only non-vanishing Poisson-bracket is

$$\{g(c), p\} = \frac{\beta}{6}\frac{\partial g}{\partial c} = \frac{\beta}{12}\sum_{\mu\in\mathbb{R}}(i\mu\xi_\mu)\exp\left(i\frac{\mu c}{2}\right), \tag{6.7}$$

which for the basic variables $h_\mu(c) := \exp\left(i\frac{\mu c}{2}\right)$ entails

$$\{h_\mu, p\} = \frac{\beta}{12} i\mu h_\mu.$$

3. Loop representation of the holonomy-flux algebra

In full LQG the topological space $\overline{\mathcal{A}}$ can equivalently be constructed as the spectrum of a certain C*-algebra. To keep the treatment simple we only presented the projective construction of this space in part 1. However, for cosmological models it is convenient to have access to a C*-algebraic description. In order to be able to better compare the present results to full LQG we therefore start by making this equivalence explicit in the realm of LQG before discussing the cosmological situation. We have to assume a certain familiarity with the theory of C*-algebras. The presentation given in the compulsory courses on functional analysis at our university should be sufficient. Readers who are unfamiliar with the theory should consult e.g. the highly recommended book by Rudin [35].

3.1. Addendum for full LQG. The equivalence of C*-algebraic approach and projective description can be found in great detail in Thiemann's book [40]. We give a simplified version of his presentation here. We will use the notation introduced in section 2 of chapter 3.

In order to give a well-defined mathematical presentation, it is necessary to make our definition of cylindrical functions more precise. For that purpose, we first define an auxiliary set $\mathrm{Cyl}'(\overline{X}) := \cup_{l\in\mathcal{L}} C(X_l)$, which is the union of all continuous, complex-valued functions on the sets X_l (cf. definition 3.7). An equivalence relation on this set is given by saying that $f \sim f'$ with $f \in C(X_l)$ and $f' \in C(X_{l'})$ iff for all l'' with $l, l' \prec l''$ the pull-backs on $X_{l''}$ with the respective projection maps agree. In mathematical notation the condition is $p_l^{l''*} f_l = p_{l'}^{l''*} f_{l'}$, where * denotes the pull-back and the maps $p_l^{l'}$ were introduced in definition 3.8. Using the fact that the set $\mathcal{L}$ is directed and (3.17), one easily sees that this condition automatically already holds for all l'' with $l, l' \prec l''$ if it holds for only one such l''. We define the cylindrical functions as the set of equivalence classes, $\mathrm{Cyl}(\overline{X}) := \mathrm{Cyl}'(\overline{X})/\sim$. Note that similar to the familiar L^p-spaces we call the equivalence classes $[f]_\sim$ cylindrical *functions*, even though they are not really functions.

The crucial observation for the definition of the C*-algebra is that for any two cylindrical functions f, f' we can find a common label $l \in \mathcal{L}$, so that both equivalence classes have a member $f_l, f'_l \in C(X_l)$ and we have $f = [f_l]_\sim$, $f' = [f_{l'}]_\sim$. The proof is almost immediate if one remembers again that the set $\mathcal{L}$ is directed.

This observation allows us to define the algebraic operations of an algebra. We define e.g. the sum of f, f' as $f + f' := [f_l + f'_l]_\sim$. The other algebraic operations can be defined

similarly. The functions which take the constant value $z \in \mathbb{C}$ everywhere can be defined by $f^z := [f_l^z]_\sim$, where $f_l^z \colon X_l \to \mathbb{C}, x_l \mapsto z$. In particular we have the function f^1 which takes the constant value 1 everywhere. Also, a norm for $f = [f_l]_\sim$ is defined by $\|f\| := \sup_{x_l \in X_l} |f_l(x_l)|$. Of course one has to verify that all these definitions are independent of the chosen representative. The respective calculations are almost immediate and pose no difficulty and are therefore left to the reader. However, we *do* want to point the reader to one very important detail, namely the fact that the compactness of the spaces X_l is crucial in the definition of the norm!

We have now defined all the necessary operations for cylindrical functions to obtain a normed, unital *-algebra. Unfortunately the corresponding metric space is not complete. This can by remedied with the standard completion procedure known from functional analysis. Since also the C*-property $\left\|f\overline{f}\right\|$ holds and extends to the completion, we obtain a unital Abelian C*-algebra $\overline{\mathrm{Cyl}(\overline{X})}$.

Let us recall two important facts from the theory of Abelian C*-algebras. First, the spectrum $\Delta\left(\overline{\mathrm{Cyl}(\overline{X})}\right)$ (the set of all homeomorphisms from the C*-algebra into the standard C*-algebra $\mathbb{C}$), which carries the Gel'fand topology (the restriction of the weak-* topology on the topological dual) is a compact Hausdorff space. Second, we recall that the Gel'fand transformation

$$\overline{\mathrm{Cyl}(\overline{X})} \to C(\Delta(\overline{\mathrm{Cyl}(\overline{X})})), f \mapsto \check{f} \text{ with } \check{f}(\chi) := \chi(f)$$

is an isometric isomorphism. Observe here that the target space of the Gel'fand transformation consists of continuous functions on a compact Hausdorff space (as is the case with any unital Abelian C*-algebra). But due to the special nature of our C*-algebra, also the source of the Gel'fand transformation consists of continuous functions on the compact Hausdorff space $\overline{X}$. This strongly hints that the spaces $\overline{X}$ and $\Delta(\overline{\mathrm{Cyl}(\overline{x})})$ are homeomorphic. Also a possible homeomorphism almost suggests itself,

$$\mathcal{X} \colon \begin{cases} \overline{X} \to \Delta(\overline{\mathrm{Cyl}(\overline{X})}) \\ x = (x_l)_{l \in \mathcal{L}} \mapsto \mathcal{X}(x), \text{ with } (\mathcal{X}(x))(f) := f_l(x_l). \end{cases}$$

Just as before, the definition is independent of the representative f_l of the cylindrical function f. The fact that this linear map is only densely defined poses no problem, since continuity can be shown by standard methods. The proof that $\mathcal{X}$ is really a homeomorphism is quite tedious and we refer the reader to [40]. Anyway, we were only interested in the general case in order to draw the analogy to the cosmological situation and will now discuss the latter in more detail.

3.2. Transferring the construction to LQC. Now that we are aware of the C*-algebraic description of the topological space $\overline{\mathcal{A}}$ in full LQG, the interpretation of almost periodic functions as cylindrical functions appears in a new light. When endowed with the usual supremum norm the functions (6.6) form a *-algebra $\mathcal{C}$. Completion in the norm then yields an Abelian C*-algebra $\overline{\mathcal{C}}$. It is therefore natural to define a distributional

extension $\overline{\mathbb{R}}$ of the configuration space $\mathbb{R}$ as the spectrum of this C*-algebra. $\overline{\mathbb{R}} := \Delta(\overline{\mathcal{C}})$ is well-known to mathematicians under the name *Bohr compactification* of the real line. Let us investigate the spectrum a bit closer in order to see how the real numbers are embedded in $\overline{\mathbb{R}}$ and also to gain some intuition on how to find a projective description of this space. Once again we roughly follow Thiemann's exposition in [40].

By definition $\overline{\mathbb{R}}$ consists of *-homomorphisms $\chi\colon \overline{\mathcal{C}} \to \mathbb{C}$. In particular χ is a linear map and can thus be defined by its values on the basis $\left\{T_\mu\colon \frac{c}{2} \mapsto \exp\left(i\mu\frac{c}{2}\right) \mid \mu \in \mathbb{R}\right\}$. Those values lead to a function X^χ via the definition $X^\chi(\mu) := \chi\left(T_\mu\right)$. The condition that χ be a *-homomorphism yields the additional constraints

$$X^\chi(\mu)X^\chi(\mu') = X^\chi(\mu+\mu') \text{ and } \overline{X^\chi(\mu)} = X^\chi(-\mu). \tag{6.8}$$

Together they give the condition $|X^\chi(\mu)|^2 = 1$, from which we read off that $X^\chi\colon \mathbb{R} \to U(1)$ is a group homomorphism. For convenience let us introduce the notation $\mathrm{Hom}\left(\mathbb{R}, U(1)\right)$ for those homomorphisms. Since the condition that X^χ is a member of this set is sufficient to guarantee that the corresponding map χ is a character of $\overline{\mathcal{C}}$, we see that $\overline{\mathbb{R}}$ can be characterized as the set $\mathrm{Hom}\left(\mathbb{R}, U(1)\right)$. The map $\overline{\mathbb{R}} \to \mathrm{Hom}(\mathbb{R}, U(1))\colon \chi \mapsto X^\chi$ is the required bijection. This interpretation is of course a perfect analogy to the definition of $\overline{\mathcal{A}}$ in the full theory contained in definition 3.3. In this context it is also useful to note that $X^\chi \in \mathrm{Hom}\left(\mathbb{R}, U(1)\right)$ need not be continuous (the values of a linear function on a basis can be defined arbitrarily) — the corresponding character $\chi\colon \overline{\mathcal{C}} \to \mathbb{C}$ in contrast *is* continuous (in fact it is a standard result of the theory of C*-algebras that any character χ is continuous and even satisfies $\|\chi\| = 1$,cf. theorem 10.7 (c) in [35]). For an explicit proof that the set $\mathrm{Hom}\left(\mathbb{R}, U(1)\right)$ even contains everywhere discontinuous elements we refer the reader to [40].

To see how $\mathbb{R}$ is embedded in $\overline{\mathbb{R}}$ we investigate once differentiable homomorphisms X. Differentiating (6.8) by μ' and setting $\mu' = 0$ we see that $X'(\mu) = X'(0)X(\mu)$. Differentiable X-s are therefore of the form $X(\mu) = e^{i\mu x} = T_\mu(x)$ with some $x \in \mathbb{R}$. We conclude that $\mathbb{R}$ is embedded in $\overline{\mathbb{R}}$ via the evaluation functionals: $x \mapsto \chi_x$ with $\chi_x(f) = f(x)$. Just as in the full theory $\overline{\mathcal{A}}$ was a distributional extension of the configuration space $\mathcal{A}$ we see that now $\overline{\mathbb{R}}$ is an extension of the configuration space $\mathbb{R}$.

These results already hint at the fact that the configuration space $\overline{\mathbb{R}}$ can also be constructed by projective methods. We try to follow the construction in the full theory (cf. section 2.1 of chapter 3) as closely as possible and outline the similarities wherever possible. Table 1 shows the corresponding concepts at a glance.

In the full theory the label set $\mathcal{L}$ of the projective family consisted of subgroupoids generated by graphs, the edges of which had to be independent in a certain sense. The graphs are now replaced by a finite set of real numbers $\gamma = \{\mu_1, \ldots, \mu_{n_l}\}$ which are

	LQG	LQC
edges e	equivalence classes of curves in Σ	real numbers μ denoting the length of an edge
graphs γ	set of independent edges	set of ILI numbers μ_i
algebraic structure of $l(\gamma)$	groupoid	group
finding γ'' with $l(\gamma), l(\gamma') \prec l(\gamma'')$	subdividing edges	choosing a basis in the submodule generated by $\gamma \cup \gamma'$
X_l	$\mathrm{Hom}(l, SU(2))$	$\mathrm{Hom}(l, U(1))$
labels for SNWF	$(e, j) \in \mathcal{P} \times \frac{1}{2}\mathbb{N}$	$\mu \in \mathbb{R}$

TABLE 1. Corresponding concepts of loop quantum gravity and -cosmology. Only concepts that actually differ are listed. The definitions of the projections $p_l^{l'}$, of cylindrical functions and of the measure μ_0 are completely equivalent.

required to be integrally linear independent.[3] We will see in a moment why this condition is necessary. Although we motivated this choice using the characterization of $\mathbb{R}$ as $\mathrm{Hom}(\mathbb{R}, U(1))$ we want to emphasize the geometrical picture. Initially the numbers μ_i just denoted the length of an edge. However, holonomies of homogeneous and isotropic connections depend only on the length of an edge and not on its location or direction. Therefore the numbers μ_i are a sufficient replacement for edges also in a geometrical sense.

These graphs generate subgroups $l(\gamma) := \{\sum_{i=1}^{n_l} k_i\mu_i \mid k_i \in \mathbb{Z}\}$ of $(\mathbb{R}, +)$ as opposed to the sub*groupoids* $l(\gamma)$ of LQG. In order to qualify as a label set of a projective family we need to make the set $\mathcal{L}$ of these subgroups into a partially ordered and directed set. We denote the generators of $l(\gamma) =: l$ by μ_i and those of $l(\gamma') =: l'$ by μ'_i and define $l \prec l'$ iff all μ_i can be expressed in the form

$$\mu_i = \sum_{j=1}^{n_{l'}} k_{ij}\mu'_j \text{ with } k_{ij} \in \mathbb{Z}. \tag{6.9}$$

The k_{ij} are then uniquely determined because γ' is an ILI set. The fact that this relation define a partial order is almost not worth mentioning. Directedness on the other hand is not as obvious. We recall that also in full LQG we could not simply take $l(\gamma \cup \gamma')$ as a common successor of l and l'. The problem arose when edges of γ and γ' intersected — in that case we had to cut those edges into pieces in order to obtain a graph consisting of independent edges. In the case now under consideration we can define γ'' as the smallest set of ILI numbers whose integral span equals the integral span of the union $\gamma \cup \gamma'$. In

[3]By integrally linear independent (ILI) we mean linear independent in the sense that we interpret the real numbers as a module over the ring $\mathbb{Z}$ of integers. In other words a linear combination $\sum_{i=1}^{n_l} k_i\mu_i = 0$ with $k_i \in \mathbb{Z}$ can only equal zero if all the coefficients vanish, $k_i = 0$. Thiemann actually uses rationally linear independent numbers, but he proves that rationally and integrally linear independent is equivalent. We prefer to use the latter notion because in our context it is the condition that we need.

other words we just choose a basis in the submodule spanned by $\gamma \cup \gamma'$. We then obviously have $l, l' \prec l''$.

The definition of the sets X_l is only slightly modified, $X_l := \mathrm{Hom}(l, U(1))$. They inherit a compact Hausdorff topology from $U(1)^{n_l}$ with the identification

$$X_l \to U(1)^{n_l} \colon x_l \mapsto (x_l(k_i))_{i \in \{1,\dots,n_l\}}.$$

Similarly, the projections $p_l^{l'}$ are just given by restricting homomorphisms $x_{l'} \colon X_{l'} \to U(1)$ to the subgroup l. Since they are obviously surjective and satisfy the consistency conditions (3.17) our definition of the projective family is complete. The proof that its projective limit $\overline{X}$, defined by (3.18), is a compact Hausdorff space can now be copied from the corresponding proof of proposition 3.11. In doing so we just have to replace the expression (3.19) of an edge e in terms of the edges of γ' by our expression (6.9).

Let us now give an explicit proof that the projective and the C*-algebraic definitions of the Bohr compactification are equivalent.

Theorem 6.1. *The map*

$$\Phi \colon \begin{cases} \overline{\mathbb{R}} & \to \overline{X} \\ \chi & \mapsto (X^\chi|_l)_{l \in \mathcal{L}}\,, \end{cases} \tag{6.10}$$

which maps a character χ to the family of restrictions of the corresponding homeomorphisms X^χ to the subgroups l, is a homeomorphism.

Proof. That $(X^\chi|_l)_{l \in \mathcal{L}}$ satisfies the consistency conditions (3.17) and is therefore contained in $\overline{X}$ is obvious. To proof injectivity suppose that $\Phi(\chi) = \Phi(\chi')$. Then $X^\chi|_l = X^{\chi'}\big|_l$ for all l and therefore also $X^\chi(\mu) = X^{\chi'}(\mu)$ for all $\mu \in \mathbb{R}$ and $\chi = \chi'$.

For surjectivity, let $x = (x_l)_{l \in \mathcal{L}} \in \overline{X}$. For every $\mu \in \mathbb{R}$ we choose an $l_\mu \in \mathcal{L}$ such that $\mu \in l_\mu$.[4] Defining $X^{\chi_x}(\mu) := x_{l_\mu}(\mu)$ our element x will then be the image $\Phi(\chi_x)$. We still have to verify that the definition of χ_x is independent of the chosen set of l_μ. Let therefore l'_μ be a different choice. Then the fact that $\mathcal{L}$ is directed guarantees the existence of l''_μ such that $l_\mu, l'_\mu \prec l''_\mu$ and $x_{l_\mu}(\mu) = x_{l''_\mu}\big|_l(\mu) = x_{l'_\mu}(\mu)$. That $\chi_x \in \overline{\mathbb{R}}$ follows from $X^{\chi_x} \in \mathrm{Hom}(\mathbb{R}, U(1))$ which in turn follows from $X^{\chi_x}|_l$ being a homomorphism for each $l \in \mathcal{L}$.

To show continuity of Φ we suppose (χ^α) is a net in $\overline{\mathbb{R}}$ converging to χ. By definition of the Gel'fand topology we then have in particular $\chi^\alpha(T_\mu) \to \chi(T_\mu)$ for all $\mu \in \mathbb{R}$ which means that

$$X^{\chi^\alpha}(\mu) \to X^\chi(\mu). \tag{6.11}$$

For continuity of Φ we have to show that $\Phi(\chi^\alpha) \to \Phi(\chi)$. By definition of the product topology this will be the case if $X^{\chi^\alpha}\big|_l \to X^\chi|_l$ for all $l \in \mathcal{L}$, where now convergence refers to the topology on X_l, which is identified with $U(1)^{n_l}$. Again by the definition of this

[4]That this can be done follows directly from the axiom of choice.

product topology we have to show that $X^{\chi^\alpha}\big|_l(\mu_i) \to X^\chi\big|_l(\mu_i)$, for all generators μ_i of l. But this statement is just a special case of (6.11).

The proof of continuity of the inverse Φ^{-1} follows the same reasoning. We assume that $X^{\chi^\alpha}\big|_l \to X^\chi\big|_l$ for all $l \in \mathcal{L}$, where we already used the product topology on $\overline{X}$. Since this applies for all l we have convergence of the $U(1)$ elements $X^{\chi^\alpha}(\mu) \to X^\chi(\mu)$ for all $\mu \in \mathbb{R}$. When we rewrite this in the form $\chi^\alpha(T_\mu) \to \chi(T_\mu)$ and use that fact that all χ^α and χ are linear we see that $\chi^\alpha(f) \to \chi(f)$ for all $f \in \mathcal{C}$. A general $f \in \overline{\mathcal{C}}$ can be approximated by f_ε such that $\|f - f_\varepsilon\| \leq \frac{\varepsilon}{3}$ for arbitrary fixed $\varepsilon > 0$. Since we have just shown convergence of the f_ε we find an index $\alpha(\varepsilon)$ so that for all α with $\alpha(\varepsilon) \prec \alpha$ we have $|\chi^\alpha(f_\varepsilon) - \chi(f_\varepsilon)| \leq \frac{e}{3}$. Combining these statements and using the triangle equality we obtain for all α with $\alpha(\varepsilon) \prec \alpha$

$$\begin{aligned}\|\chi^\alpha(f) - \chi(f)\| &\leq |\chi^\alpha(f - f_\varepsilon)| + |\chi(f - f_\varepsilon)| + |\chi^\alpha(f_\varepsilon) - \chi(f_\varepsilon)| \\ &\leq 2\,\|f - f_\varepsilon\| + \frac{\varepsilon}{3} \leq \varepsilon,\end{aligned}$$

where we made use of the fact that as characters χ and χ^α are bounded by 1. This establishes convergence $\chi^\alpha(f) \to \chi(f)$ for all $f \in \overline{\mathcal{C}}$. By definition of the Gel'fand topology then $\chi^\alpha \to \chi$. □

This theorem completes our detailed investigation of the space $\overline{\mathbb{R}}$. As in full LQG we want to use a Hilbert space of the form $L^2(\overline{\mathbb{R}}, d\mu_0)$ to define our representation. Thus we have to construct the measure μ_0. The construction is entirely parallel to that of the Ashtekar measure in full LQG, so we can keep the discussion short. In LQG we defined the measure via the Riesz representation theorem and a corresponding linear functional Λ_0 on the space $C(\overline{\mathcal{A}})$. This functional was defined in (3.21) by its values on the SNWF, which are linearly independent and whose linear combinations are dense in $C(\overline{\mathcal{A}})$. In LQC the role of SNWF is played by the Gel'fand transforms $\check{T}_\mu$, defined by $\check{T}_\mu(\chi) := \chi(T_\mu)$. The natural generalization of (3.21) is therefore $\Lambda_0(\check{T}_\mu) := \delta_{\mu,0}$. Observe that from the theorem of Gel'fand Neimark (which guarantees that the Gel'fand transformation is an isometric isomorphism) we know that Λ_0 is densely defined on a linearly independent set. Furthermore this powerful theorem establishes

$$\Lambda_0(\overline{\check{T}_\mu}\check{T}_{\mu'}) = \Lambda_0(\check{T}_{\mu'-\mu}) = \delta_{\mu,\mu'}, \tag{6.12}$$

showing that Λ_0 is a positive linear functional. As representation space we can therefore take $L^2(\overline{\mathbb{R}}, d\mu_0)$, where μ_0 is the regular Borel probability measure that represents Λ_0 in the sense of the Riesz representation theorem D.4. We mention that (6.12) also establishes orthonormality of the Gel'fand transforms $\check{T}_\mu$. Since it is easily seen that all the premises of the Stone-Weierstrass theorem D.3 are met, this set is also dense in $C(\overline{\mathbb{R}})$. The set of Gel'fand transforms $\check{T}_\mu$ is therefore an orthonormal basis of our representation space, which perfectly fits with LQG, where the SNWF also constitute an orthonormal basis.

Now that we have defined the Hilbert space $L^2(\overline{\mathbb{R}}, d\mu_0)$, we use it to define a representation of the algebra (6.7). Just as in full LQG (cf. section 2.3 in chapter 3) we let

symmetric cylindrical functions $g \in \mathrm{Cyl}_S$ act as multiplication operators,

$$(\widehat{g}f)(\chi) := g(\chi)f(\chi), \quad \forall f \in L^2(\overline{\mathbb{R}}, d\mu_0).$$

The momenta p act as derivative operators

$$\widehat{p} := -\frac{i\beta}{6}\frac{d}{dc}$$

and are a priori only defined on the dense subspace Cyl_S (or its image under the Gel'fand transform to be precise).

Let us introduce the bra-ket notation for the basis states and write $\langle c|\mu\rangle := T_\mu(c)$. The action of the momentum operators on these states is

$$\widehat{p}\,|\mu\rangle = \frac{\beta\mu}{12}\,|\mu\rangle =: p_\mu\,|\mu\rangle\,.$$

Finally we mention that in this new approach to LQC the volume is quantized directly in the form (5.14). The states $|\mu\rangle$ are therefore also eigenstates of the volume operator with eigenvalues

$$V_\mu = |p_\mu|^{\frac{3}{2}} = \left(\frac{\beta\,|\mu|}{12}\right)^{\frac{3}{2}} l_P^3. \tag{6.13}$$

4. Quantization of the Hamiltonian constraint

In order to see the connection to the traditional formalism presented in the previous chapter we now discuss dynamics in the new setting. We start with the expression (5.23) of the Hamiltonian constraint in terms of a Poisson bracket with the volume operator. In order to regularize this constraint we are now in a better position since we are not working with point holonomies but with actual holonomies. We can therefore omit the complicated shrinking of the auxiliary manifold and directly work on the spatial manifold Σ. On the other hand our situation is also worse than it was before. In working with matrix elements of holonomies in our definition (6.6) we implicitly dispensed with all information that refers to directions in the spatial manifold. When we now want to approximate the curvature components F^i_{ab} by the holonomy along a square we have to reintroduce some of this information.

The way to incorporate directional structure is to use holonomies along edges tangent to the fiducial cell $\mathcal{V}$ introduced above. In other words the edges $e_i^{(\mu_0)}$ are parallel to the vectors X_i^a and have length $\mu_0 V_0^{\frac{1}{3}}$ with respect to $\mathring{q}_{ab}$. We define

$$h_i^{(\mu_0)} := \mathcal{P}\exp\left(\int_{e_i^{(\mu_0)}} A_a^j \tau_j\right) = \exp\left(\mu_0 c \tau_i\right). \tag{6.14}$$

Two of these edges span a square $\alpha_{ij} := (e_j^{(\mu_0)})^{-1} \circ (e_i^{(\mu_0)})^{-1} \circ e_j^{(\mu_0)} \circ e_i^{(\mu_0)}$ which can be used to approximate the curvature components. The holonomy along the square is

$$h_{\alpha_{ij}}^{(\mu_0)} = \mathbb{1} - \mu_0^2 F^k_{ab}\tau_k X_i^a X_j^b + O(\mu_0^3), \tag{6.15}$$

as can be verified with the expansion

$$h_i^{(\mu_0)} = \mathbb{1} + \mu_0 c\tau_i + \frac{1}{2}\mu_0^2 c^2 \tau_i^2 + O(\mu_0^3) \tag{6.16}$$

by a calculation completely equivalent to (5.26) (just replace ε by μ_0 and Φ_I by $c\tau_i$).

With the help of the last two equations we can now verify that for

$$H^{(\mu_0)}(N) := -\frac{8N}{\beta^3\mu_0^3}\sum_{ijk}\varepsilon^{ijk}\,\mathrm{tr}\left((h_j^{(\mu_0)})^{-1}(h_i^{(\mu_0)})^{-1}h_j^{(\mu_0)}h_i^{(\mu_0)}(h_k^{(\mu_0)})^{-1}\left\{h_k^{(\mu_0)}, V\right\}\right), \tag{6.17}$$

the limit $\lim_{\mu_0\to 0} H^{(\mu_0)}(N) = H(N)$ reproduces the Hamiltonian constraint in the second form in (6.4). The proof of this statement follows the same reasoning as the ones for the corresponding statements in full LQG and in traditional LQC. It can therefore be taken without noticeable modifications from the proof of (5.27).

In passing to quantum theory we promote holonomies to multiplication operators and replace the Poisson bracket by a commutator. With $h_i^{(\mu_0)} = \cos\frac{\mu_0 c}{2} + 2\sin\frac{\mu_0 c}{2}\tau_i$ we arrive at

$$\begin{aligned} H^{(\mu_0)}(1) &:= \frac{8Ni}{\beta^3\mu_0^3}\sum_{ijk}\varepsilon^{ijk}\,\mathrm{tr}\left(\widehat{h}_i^{(\mu_0)}\widehat{h}_j^{(\mu_0)}(\widehat{h}_i^{(\mu_0)})^{-1}(\widehat{h}_j^{(\mu_0)})^{-1}(\widehat{h}_k^{(\mu_0)})^{-1}\left[\widehat{h}_k^{(\mu_0)}, V\right]\right) \\ &= \frac{192i}{(\beta\mu_0)^3}\sin^2\frac{\mu_0 c}{2}\cos^2\frac{\mu_0 c}{2}\left(\sin\frac{\mu_0 c}{2}\widehat{V}\cos\frac{\mu_0 c}{2} - \cos\frac{\mu_0 c}{2}\widehat{V}\sin\frac{\mu_0 c}{2}\right). \end{aligned}$$

The calculations that lead to this expression are entirely parallel to the ones that lead to (5.31) (just replace c by $\mu_0 c$ and Λ_I by τ_i).

To calculate the action on the basis consisting of the states $|\mu\rangle$ we need the action of the operators $\cos\frac{\mu_0 c}{2}$ and $\sin\frac{\mu_0 c}{2}$. Using (5.34) we obtain

$$\begin{aligned} \cos\frac{\mu_0 c}{2}|\mu\rangle &= \frac{1}{2}\left(|\mu+\mu_0\rangle + |\mu-\mu_0\rangle\right), \\ \sin\frac{\mu_0 c}{2}|\mu\rangle &= \frac{1}{2i}\left(|\mu+\mu_0\rangle - |\mu-\mu_0\rangle\right), \end{aligned}$$

which precisely mirrors (5.35). Since the states μ are also eigenstates of the volume operator with eigenvalues (6.13) the calculation of the action of $\widehat{H}^{(\mu_0)}(1)$ on $|\mu\rangle$ is again equivalent to that in traditional LQC. We therefore have

$$\widehat{H}^{(\mu_0)}(1)|\mu\rangle = \frac{6}{(\beta\mu_0)^3}(V_{\mu+\mu_0} - V_{\mu-\mu_0})\left(|\mu+4\mu_0\rangle - 2|\mu\rangle + |\mu-4\mu_0\rangle\right). \tag{6.18}$$

First of all note the similarity with (5.36) (cf. also the next chapter). We observe that our action of the Hamiltonian constraint still contains the regularization parameter μ_0. Unfortunately we can not take the limit $\mu_0 \to 0$ because the limit of the classical expression (6.17) contains c^2 and there is no operator corresponding to this phase-space function. We are facing the same problem as in traditional LQC. The diffeomorphism constraint is already solved, so we can not use the same procedure as in the full theory to obtain a regulator-independent expression.

The viewpoint regarding the dependence of the Hamiltonian constraint on μ_0 is inspired by a similar situation in full LQG. There we also had to introduce an auxiliary edge to approximate the curvature components (in figure 7 this was the edge a_{23}). This auxiliary edge is labeled by a spin j. Diffeomorphism invariance ensures that the Hamiltonian constraint does not depend on the edge e, but it *does* depend on the spin j. There is a quantization ambiguity and a way to fix it is to assign the lowest possible non-trivial value $j = \frac{1}{2}$.

This analogy suggests that we regard $\widehat{H}^{(\mu_0)}$ as an exact constraint and call μ_0 a quantization ambiguity. Unfortunately there is no smallest real number $\mu_0 > 0$ which we could use to fix the ambiguity. However, in LQG there is also a more geometrical motivation for choosing $j = \frac{1}{2}$ which *can* be transcribed to our problem. Remember that the auxiliary edge was introduced to approximate the curvature components F^i_{ab}. Classically one would use a loop enclosing an infinitesimally small area. However, we have seen that this is not possible in quantum geometry. At best we could use a loop enclosing an area $A_0 = \frac{\sqrt{3}\beta l_P^2}{4}$, which is the smallest possible eigenvalue of the area operator according to (3.31). Transferring these ideas to our situation we are lead to setting the value of μ_0 such that the area spanned by α_{ij} equals A_0. Now the expression (6.13) for the volume operator suggests that we can regard $\widehat{|p|}$ as quantization of the area operator. The auxiliary edge $h_k^{(\mu_0)}$ defines a state $|\mu_0\rangle$ and the eigenvalue of the area operator in this state is p_μ. The idea is now to set this eigenvalue, which represents the area of the surface spanned by α_{ij}, to A_0. In other words we set $\frac{\beta\mu_0}{12} = A_0 = \frac{\beta}{2}\frac{\sqrt{3}}{4}$, which means $\mu_0 := \frac{3\sqrt{3}}{2}$.[5]

[5]We remark that there was a typographical mistake for this value in the original article [5], which was corrected in [7].

CHAPTER 7

Discussion and Comparison

In this last chapter we first present some results of LQC that are direct consequences of the results in the previous chapters and are physically important. Thereafter we compare the two different approaches to the kinematics of LQC. We will see that both have their respective merits and drawbacks. The last section discusses various suggestions for constructing a theory that unites the advantages of both approaches.

1. Interesting results of loop quantum cosmology

The results mentioned now all hold for both traditional LQC and the new approach to kinematics. When we give any details we present them in the latter formalism and follow [5]. The detailed calculations for traditional LQC were carried out in [15, 14, 12].

1.1. Boundedness of the inverse scale factor. One of the strongest results of LQC, which is already obtained at the kinematical level, is that the inverse scale factor $a^{-1} = |p|^{-\frac{1}{2}}$ is bounded above. Remembering that in our reduced classical models curvature is a multiple of a^{-2} and that any matter Hamiltonian involves a term containing a power of a^{-1} this is a very powerful result. It means that the classical curvature singularity at the big bang is naturally resolved by our quantization method. Let us briefly mention how to arrive at this important statement, so that one does not get the impression that it simply arises by some ad hoc assumption.

The obvious quantization of any function of p would be via the spectral theorem. Since all eigenvectors of $\widehat{p}$ are normalizable ($\langle \mu_1, \mu_2 \rangle = \delta_{\mu_1,\mu_2}$) the decomposition of the identity is a direct sum, not an integral, $\mathbb{1} = \sum_\mu |\mu\rangle \langle\mu|$. Now the operator $\widehat{p}$ has an eigenvector $|0\rangle$ with eigenvalue 0. Therefore the definition of $\widehat{p^{-1}}$ via the spectral theorem yields a division by zero and is therefore inapplicable. Hence we are forced to use a different quantization of the inverse scale factor and are naturally led to the strategy already employed for the quantization of the Hamiltonian constraint in full and reduced theory. One expresses $\operatorname{sgn} p\, |p|^{-\frac{1}{2}}$ as a Poisson bracket $\sum_i \tau^i h_i \left\{ h_i^{-1}, V^{\frac{1}{3}} \right\}$ of a holonomy with the volume operator.[1] The corresponding operator, which has holonomies replaced by multiplication operators and Poisson brackets by commutators has the states $|\mu\rangle$ as eigenvectors with eigenvalues

$$\frac{12}{\beta l_P^2} \left(V_{\mu+1}^{\frac{1}{3}} - V_{\mu-1}^{\frac{1}{3}} \right).$$

[1] We mention that in contrast to the quantization of the Hamiltonian constraint this expression is exact, i.e. it does not depend on a parameter μ_0. The complicated question on how to deal with a regulator whose limit we cannot take therefore does not arise here.

Inserting the values (6.13) for V_μ one quickly checks that our operator is bounded above. The upper bound is obtained for $\mu = 1$ and equals $\sqrt{\frac{24}{\beta}} l_P^{-1}$. The curvature therefore has an upper bound proportional to $\frac{24}{\beta l_P^2}$. Observe that in the classical limit $\hbar \to 0$ this bound tends to infinity, just as one would expect.

Summarizing, we can say that the classical curvature singularity at the big bang is resolved by LQC and this result does not come about because of the introduction of some cut off. Instead it arises naturally by following the methods used for full LQG.

1.2. Discreteness of time-evolution and classical limit. As was the case in full LQG, the Hamiltonian constraint (6.18) has no solutions in the kinematical Hilbert space. We have to use a Gel'fand triple

$$\mathrm{Cyl}_S \hookrightarrow \mathcal{H}_{\mathrm{kin}} \hookrightarrow \mathrm{Cyl}_S^*$$

and solutions will be distributions in Cyl_S^*. We expand those elements in the basis provided by the eigenstates of $\widehat{p}$,[2]

$$(\Psi| = \sum_\mu \psi(\mu) \langle \mu| \,. \tag{7.1}$$

The Hamiltonian constraint $(\Psi|\big(\widehat{H}^{(\mu_0)}(1)\,|\nu\rangle\big) = 0$ for all $\nu \in \mathbb{R}$ then yields the difference equation,[3]

$$\begin{aligned}(V_{\mu+5\mu_0} - V_{\mu+3\mu_0})\,\psi(\mu+4\mu_0) - 2\,(V_{\mu+\mu_0} - V_{\mu-\mu_0})\,\psi(\mu)\\ + (V_{\mu-3\mu_0} - V_{\mu-5\mu_0})\,\psi(\mu-4\mu_0) = 0\end{aligned} \tag{7.2}$$

which constrains the coefficients $\psi(\mu)$. One can fix two of the coefficients $\psi(\mu)$, $\psi(\mu+4\mu_0)$ and the Hamiltonian constraint determines all values of $\psi(\mu+4\mu_0 n)$ for $n \in \mathbb{N}$.

Let us outline some important aspects of the dynamics. First, we want to stress the importance of the fact that the Hamiltonian constraint turns into a *difference* equation rather than a *differential* equation. We therefore have a discrete time-evolution, a feature that is expected from any quantum theory of gravity. The second important feature is that one can evolve through the classical singularity at $\mu = 0$. A detailed investigation of the difference equation shows that the coefficient $\psi(0)$ remains undetermined and therefore decouples from the rest.[4] This fact has coined the expression "big bounce", which replaces the classical big bang in LQC. We also want to point the reader to the interesting detail that the question of what happens to the coefficient $\psi(0)$ is only relevant if we attack the singularity directly, i.e. we start with coefficients $\psi(-4\mu_0 N)$ and $\psi(-4\mu_0(N-1))$ for some large, positive integer N. If we take different starting points for the value of ψ, the

[2]Elements of the algebraic dual space can have uncountably many non-vanishing coefficients $\psi(\mu)$. We denote these elements by $(\Psi|$ instead of $\langle\Psi|$, so that we don't confuse them with elements of $\mathcal{H}_{\mathrm{kin}}$.

[3]In a physically realistic scenario one would have to include a matter contribution on the RHS and make the coefficients ψ depend on the matter field.

[4]In a more realistic model one has to include a matter term in the Hamiltonian. It turns out that for all minimally coupled matter Hamiltonians this statement remains true since they annihilate $\psi(0)$.

time-evolution does not even take notice of what happens at $\mu = 0$, it jumps right over this value of the classical singularity.

Finally let us briefly comment on some important consistency conditions fulfilled by LQC. It is shown in [5] that in the limit $\mu_0 \to 0$ (which contradicts the ideas of quantum geometry but is mathematically interesting nonetheless) the Hamiltonian constraint reduces to the familiar Wheeler DeWitt equation in a precise sense. Also solutions to the Wheeler DeWitt equation — which exhibit the singular behavior of the classical theory for small value of the momentum variable p — can be regarded as approximate solutions to the difference equation (7.2) for $p \gg p_{\mu_0}$. Furthermore, by using coherent states peaked around points of the classical phase space (large μ), the expectation value of the constraint in these states is equal to the classical constraint (modulo some negligible corrections). In this sense the theory also has the correct classical limit.

2. Comparison of both approaches

Having presented two different approaches to the kinematical structure of loop quantum cosmology, the question arises which one is "correct". Here it is a priori not clear what we mean by "correct". Because of the difficulties briefly mentioned in chapter 2 we are, at least at the moment, not able to construct cosmological solutions of the full theory. The ultimate criterion of checking if the solutions of the theory under consideration correspond to symmetric[5] solutions of the full theory is therefore out of reach. We have already mentioned that the states of traditional LQC can be interpreted as distributional and in a certain sense symmetric states of LQG. But also in the new approach the solutions are contained in Cyl^*_S which is a subspace of Cyl^*. So this criterion can also not be used to rule one of our approaches out. We therefore have to take formal aspects into consideration and start with the discussion of Bojowald's original approach.

The primary criticism brought forward against Bojowald's point holonomies is that they don't separate the points of the configuration space, which is in harsh contrast to our requirement on the quantization procedure (cf. item 1 of section 1 in chapter 2). Indeed, being periodic functions of the configuration space variable c, two points which are exactly a period length apart can never be separated.[6] This is not to say that this theory can not yield correct results. Indeed, it also has desirable features not found in the new approach. The kinematical Hilbert space has the countable basis of theorem 5.2 and is thus seen to be separable. Another feature that is desirable for all symmetry reduced physical models should also be mentioned. Bojowald quantizes the more general homogeneous models and only incorporates isotropy after quantization. In this way, compared to the new approach, more degrees of freedom are quantized. Finally we mention that the procedure of removing the regulator in the Hamiltonian constraint operator of LQG is nicely mirrored. In full LQG the action on diffeomorphism invariant states was independent of the regulator; in

[5]It is not even fully clear how to define symmetric states in full LQG.

[6]As this argument shows, the point holonomies would separate the points on the compactification S^1 of the configuration space $\mathbb{R}$.

LQC the action of the Hamiltonian constraint is independent of the shrinking manifold $\overline{S/F}_\varepsilon$.

In the new approach one deals with the non-separability of the point holonomies by using *almost* periodic functions instead of periodic ones. However, this introduces new limitations. First, by quantizing directly the isotropic models all information on directions in space is thrown out right at the start. Later, when we need this information in order to approximate the curvature components appearing in the Hamiltonian constraint, we artificially have to reintroduce some of this structure. Another disadvantage lies of course in the choice of the regulator μ_0. The fact that we cannot fix μ_0 within the reduced theory is rather disappointing. Also our prescription for the choice of this parameter is not undisputed. In [8] a different prescription is presented in which μ_0 becomes a function of p.

Our most serious objection lies in the fact that we obtain a huge non-separable kinematical Hilbert space — after all in full LQG we attached rather great importance to the separability of $\mathcal{H}_{\text{diff}}$. Intuitively in a reduced theory with less degrees of freedom this space should be smaller than in full LQG. This non-separable Hilbert space also leads to a large space of solutions to the Hamiltonian constraint. We discussed briefly in the first section of this chapter that the difference equation only constrains the values of $\psi(\mu + 4\mu_0 n)$. Thus there are as many solutions as there are functions on an interval $[\mu' - 4\mu_0, \mu' + 4\mu_0)$ with an arbitrary fixed $\mu' \in \mathbb{R}$. However, in order to complete the quantization procedure one still has to introduce a physical inner product on the space of solutions to the Hamiltonian constraint. One then requires physical states to have finite norm, which eliminates a lot of redundant states. The authors of [5] claim that in simple examples this strategy suffices to regain a physical Hilbert space of reasonable size.

After all those objections let us mention that the primary appeal of the new approach lies in the rigorous construction of the representation space via C*-algebra techniques and projective methods. Even though the corresponding holonomies are valued in $U(1)$ as opposed to $SU(2)$, it is surprising how many details of the construction in full LQG can be transferred to the reduced theory (cf. section 3 of the previous chapter).

Finally, let us also mention that all the differences between the two approaches are at the kinematical level. The dynamics are actually equivalent as can be seen from the almost identical expressions of the respective Hamiltonian constraints (5.36) and (6.18).

Summarizing we can say that both approaches have their respective pros, but also have serious drawbacks. This raises the question if one could construct a theory which incorporates the merits of both theories. In the following section we will present some ideas in that direction.

3. Suggestions for improvements of the kinematical structure

As Bojowald's point holonomies do not separate the points of the configuration space, they do not make for a good starting point. We will thus start with the representation of

chapter 6 and our challenge is to find a way to obtain a separable diffeomorphism-invariant Hilbert space.

3.1. The original idea. Let us start with an alternative, simplified derivation of the form (6.1) of the canonical variables. This will provide a better understanding of the rescalings introduced in (6.2) which are at the basis of our idea.

It is well-known (see e.g. [42]) that by choosing appropriate coordinates (t, x, y, z), the metric of a homogeneous, isotropic space-time can be written in the form

$$(ds^2)_{ab} = -(dt^2)_{ab} + a^2(t)(dx^2 + dy^2 + dz^2)_{ab} =: -(dt^2)_{ab} + a^2(t)\mathring{q}_{ab}. \tag{7.3}$$

We choose a cotriad ω^i_a compatible with $\mathring{q}_{ab}$ (i.e. $\omega^i_a\omega^i_b = \mathring{q}_{ab}$) and denote the corresponding triad by X^a_i. This form of the metric provides us with a natural 3+1 split for the ADM-decomposition. A physical cotriad of the spatial metric can immediately be read off, $e^i_a = a(t)\omega^i_a$. With $\omega_q = a^3(t)\omega_{\mathring{q}}$ we have the densitized triad $\widetilde{E}^a_i = a^2(t)X^a_i\omega_{\mathring{q}}$.

The connection $A^i_a = \Gamma^i_a + \beta K^i_a$ can be found be noticing that for our flat spatial metric $\Gamma^i_a = 0$ and

$$K_{ab} = \frac{1}{2}\mathcal{L}_n h_{ab} = \frac{1}{2}\frac{d}{dt}a^2(t)\mathring{q}_{ab} = a(t)\dot{a}(t)\mathring{q}_{ab}.$$

This yields

$$A^m_a = \beta K_{ab}E^{a,m} = \beta a(t)\dot{a}(t)\mathring{q}_{ab}a^{-1}(t)X^a_m = \beta\dot{a}(t)\omega^m_b.$$

Summarizing, we have

$$A^i_a = \beta\dot{a}\omega^i_a, \text{ and } \widetilde{E}^a_i = a^2X^a_i\omega_{\mathring{q}}, \tag{7.4}$$

which is precisely our previous form (6.1) when we set $\beta\dot{a} =: \tilde{c}$ and $a^2 =: \tilde{p}$.

Now we remember our discussion of diffeomorphisms in symmetry-reduced theories from section 3 in chapter 5 — they correspond to the remaining freedom in choosing coordinates which keep the form of the invariant metric. In our case the metric (7.3) allows a change of coordinates

$$r_\lambda \colon x^i \mapsto \lambda x^i, \tag{7.5}$$

where $\lambda \in \mathbb{R} \setminus \{0\}$. This transformation maps $\mathring{q}_{ab}$ to $\lambda^2\mathring{q}_{ab}$. The form (7.3) is preserved with the new scale-factor $\frac{a^2(t)}{\lambda^2}$. We can thus find the expression for our basic variables A^i_a and $\widetilde{E}^a_i$ in the transformed coordinates by replacing $a(t)$ by $\frac{a(t)}{\lambda}$ in (7.4). We again find the same behavior under rescalings as when we first discussed them in section 1 of the previous chapter.

This discussion establishes the correspondence between rescalings of the basic variables and the remaining diffeomorphisms of LQC. The idea is now to use the rescalings (7.5) to implement a "diffeomorphism"-constraint, which is expected to reduce the size of the diffeomorphism-invariant Hilbert space. Now we have seen in chapter 5 that the diffeomorphism constraint is already solved in LQC. This can also be seen directly in the new approach to kinematics,

$$\widetilde{V}_a = \widetilde{E}^b_i F^i_{ba} = p\omega_{\mathring{q}}X^b_i c^2\varepsilon^i{}_{jk}\omega^j_b\omega^k_a = pc^2\omega_{\mathring{q}}\omega^k_a\delta^j_i\varepsilon^i{}_{jk} = 0.$$

Let us spell out the reason why our idea has nevertheless a chance of reducing the size of the diffeomorphism-invariant Hilbert space. The diffeomorphism constraint in its form (2.25) only implements invariance under *local* diffeomorphisms (this entered the derivation in section 3.2 of chapter 2 via the assumption that $\xi \lrcorner A$ has compact support so that we could drop boundary terms in the partial integration). In contrast, the rescalings r_λ are global diffeomorphisms and the corresponding constraint is not already solved. To show this explicitly, we use the method of sections 3.1 and 3.2, chapter 2 to derive the constraint. Note that in order to obtain a transformation of our basic variables we now have to regard the rescalings as an active transformation and not as a mere change of coordinates. The infinitesimal transformation $x^i \mapsto x^i + (\lambda - 1)x^i$ then causes the infinitesimal transformation $A^i_a \mapsto A^i_a + \delta A^i_a = A^i_a + (\lambda - 1)A^i_a$. This leads to the constraint

$$R(\lambda) := \int \widetilde{E}^a_i \delta A^i_a = \tilde{c}\tilde{p} \int \omega_{\hat{q}} X^a_i (\lambda - 1)\omega^i_a = 3\tilde{c}\tilde{p}V_0 = 3cp. \tag{7.6}$$

The fact that we implement invariance under diffeomorphisms has motivated our choice for the title of this thesis. Nevertheless, in order not to cause confusion with the existing diffeomorphism-constraint, from now on we will use the term rescalings-constraint.

When we now try to find solutions we use the notation $\mathcal{H}_{\text{kin}}$ for the existing kinematical Hilbert space $L^2(\overline{\mathbb{R}}, d\mu_0)$ and denote our desired space of solutions by $\mathcal{H}_{\text{diff}}$. The constraint (7.6) contains c, which has no direct operator analog in quantum theory. We therefore follow the same procedure as we did in solving the diffeomorphism constraint of full LQG (section 3.2 in chapter 3) and first define a representation of the rescalings on $\mathcal{H}_{\text{kin}}$. The transformation of the connection variable ($r_\lambda(c) = \lambda c$) suggests the definition $\widehat{R}(\lambda) \exp\left(i\mu\frac{c}{2}\right) := \exp\left(i\mu\frac{\lambda c}{2}\right)$. Writing this in the bra-ket notation we have

$$\widehat{R}(\lambda) \left|\mu\right\rangle := \left|\lambda\mu\right\rangle, \tag{7.7}$$

which defines the representation densely on a basis. The representation property is immediate and unitarity of the representation is obvious from $\langle\mu|\nu\rangle = \delta_{\mu,\nu}$.

Again there are no solutions to the constraint in $\mathcal{H}_{\text{kin}}$. To see this, suppose $f \in \mathcal{H}_{\text{kin}}$. Expanding f in our basis we have $|f\rangle = \sum_{\mu\in\mathbb{R}} \xi_\mu |\mu\rangle$ where only a countable number of the ξ_μ can be non-vanishing. The condition $\widehat{R}(\lambda)f = f$ requires $\sum_{\mu\in\mathbb{R}} \xi_\mu |\lambda\mu\rangle = \sum_{\mu\in\mathbb{R}} |\mu\rangle$ for all $\lambda \in \mathbb{R}$. Therefore, if only one ξ_μ with $\mu \neq 0$ is non-vanishing, all other ξ_μ except ξ_0 would have to take the same numerical value. This state is clearly not in $\mathcal{H}_{\text{kin}}$ and the only solution is the vacuum solution with zero volume, $|f\rangle = |0\rangle$.

We are forced to use a Gel'fand triple

$$\text{Cyl}_S \hookrightarrow \mathcal{H}_{\text{kin}} \hookrightarrow \text{Cyl}^*_S,$$

where we use the symmetric cylindrical functions (6.6) as a dense subspace of the kinematical Hilbert space. Observe that in order to find solutions to the Hamiltonian constraint one uses the same Gel'fand triple. Solutions to all the constraints will therefore be elements of Cyl^*_S which satisfy the difference equation (7.2), as well as the rescalings

constraint. Starting with the rescalings constraint we expand an element $(\Psi|$ in the form

$$(\Psi| = \sum_{\mu} \psi(\mu) \langle \mu| \,.$$

We require $(\Psi\,|f\rangle = (\Psi|\big(\widehat{R}(\lambda)\,|f\rangle\,\big)$ for all $\lambda \in \mathbb{R}\setminus\{0\}$ and all $|f\rangle \in \mathrm{Cyl}_S$. In particular for $|f\rangle = |\nu\rangle$ we have the condition $\psi(\nu) = \psi(\lambda\nu)$ for all $\lambda \in \mathbb{R}\setminus\{0\}$. Since any real number except 0 can be mapped to any other such real number by a rescaling, this can only be fulfilled by

$$\psi(\mu) = \begin{cases} \alpha_1, & \mu = 0 \\ \alpha_2, & \mu \in \mathbb{R}\setminus\{0\} \end{cases} \tag{7.8}$$

with constants $\alpha_1, \alpha_2 \in \mathbb{C}$. Unfortunately this two-dimensional space of solutions is much smaller than the separable, yet infinite dimensional space we were after. The problem is even more apparent when one takes a look at the Hamiltonian constraint (7.2). As long as $\alpha_2 \neq 0$ our functions (7.8) do not satisfy the difference equation except for $\mu_0 = 0$, which is not compatible with the ideas of quantum gravity. The only normalized solution to all the constraints would be the functional $\langle 0|$, which corresponds to a state with zero volume, i.e. an empty universe.

We want to mention that our solutions can also be obtained via a rigging map η as introduced in section 3.2, chapter 3. For a symmetric cylindrical function $f = \sum_{\mu\in I} \xi_\mu\,|\mu\rangle$ with finite index set $I \subset \mathbb{R}$ the action of the functional $\eta(f)$ is defined via

$$\eta(f)\,|\nu\rangle = \sum_{f'\in \mathrm{Orb}(f)} \langle f'|\nu\rangle = \sum_{\lambda\in\mathbb{R}\setminus\{0\}} \sum_{\mu\in I} \overline{\xi_\mu}\,\langle \lambda\mu|\nu\rangle = \begin{cases} \overline{\xi_0}, & \nu = 0 \\ \sum_{\mu\in I\setminus\{0\}} \overline{\xi_\mu}, & \nu \neq 0. \end{cases}$$

With $\overline{\xi_0} =: \alpha_1$ and $\sum_{\mu\in I\setminus\{0\}} \overline{\xi_m} =: \alpha_2$ those are precisely our solutions (7.8). The rigging map also defines an inner product on the space of solutions via $\langle \eta(f)|\eta(g)\rangle := \eta(f)\,|g\rangle$. For example we see that all solutions $\eta(|\mu\rangle)$ are normalized,

$$\langle \eta|\mu\rangle, \eta|\mu\rangle\rangle = \eta|\mu\rangle(|\mu\rangle) = \sum_{\lambda\in\mathbb{R}\setminus\{0\}} \langle \lambda\mu|\mu\rangle = 1.$$

3.2. Suggested modifications of the original idea. Since the first attempt failed to create a separable but still infinite dimensional diffeomorphism invariant Hilbert space we tried to make modifications at various stages.

One of the simplest modifications that comes to mind is only demanding invariance under rescalings r_λ with $\lambda \in \mathbb{R}^+$. Of course such a minor modification of the input does not have a chance of creating a big impact on the result. Indeed, by following the same steps as before, we see that now the solutions are parameterized by three constants $\alpha_1, \alpha_2, \alpha_3$, since the functionals can now take different values on states $|\mu\rangle$ with positive and negative μ, respectively.

Going one step further, one might also try a different Gel'fand triple

$$\mathcal{D} \hookrightarrow \mathcal{H}_{\mathrm{kin}} \hookrightarrow \mathcal{D}^*.$$

Since our original space of solutions was too small we should try to make $\mathcal{D}^*$ larger — equivalently $\mathcal{D}$ must be a subspace of Cyl_S. On the other hand $\mathcal{D}$ must still be dense in $\mathcal{H}_{\mathrm{kin}}$ and this condition is hard to match for a subspace of Cyl_S. A possible choice that might work is therefore a space $\mathcal{D}$ disjoint from Cyl_S. This briefly inspired us to try $\mathcal{D} := \left\{ f \in \mathcal{H}_{\mathrm{kin}} \mid f = \sum_{i \in I} \xi_i \left| \mu_i \right\rangle \right\}$ with a countable index set I and a genuinely infinite number of non-vanishing coefficients ξ_i (i.e. f is *not* in $\mathcal{D}$ if it only has finitely many non-vanishing ξ_i). Unfortunately this set is not even a vector space (e.g. it does not contain the zero vector), so this approach had to be discarded.

Our remaining proposals all have to do with incorporating certain features of full LQG, which are not found in the existing treatments of LQC. The first in this list is inspired by Rovelli's idea of relaxing the diffeomorphism group to the group Diff^* of continuous bijections which may fail to be analytic at a finite number of isolated points. When we try to apply this idea for our constraint we are led to consider what we want to call piecewise constant rescalings. To be precise, our scale factor is now a simple (in the sense used in integration theory) function $\lambda \colon \mathbb{R} \to \mathbb{C}$, which can be written as a linear combination of characteristic functions, $\lambda = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$. We require that $\alpha_i \neq 0, \forall i \in I$ and that the disjoint union of the intervals A_i is all of $\mathbb{R}$. Intuitively, those functions are constant on intervals of $\mathbb{R}$. But although this choice of symmetry group has a good motivation from full LQG, it has no chance of meeting our needs. The reason is that by enlarging the group of allowed rescalings the space of solution can only become smaller. By exactly the same calculations as for the original idea one quickly sees that our space of solutions is actually not reduced, we obtain the same two-dimensional space as before.[7]

The better way is therefore not to enlarge the group of rescalings, but to impose invariance under only a *sub*group. Being labeled by a real number λ, we have to find a subgroup V of the multiplicative group $(\mathbb{R} \setminus \{0\}, \cdot)$. Two natural choices for V come to mind. The first is $V_1 := (\mathbb{Q} \setminus \{0\}, \cdot)$. The only other natural choice is the group generated by the one distinguished real number we have, $V_2 := \langle \mu_0 \rangle$. Let us present an argument which shows that already the group V_1 is too small to significantly reduce the size of $\mathcal{H}_{\mathrm{diff}}$. The solutions to the constraint can be obtained in the same manner as before and can be described by arbitrary functions $\psi \colon \mathbb{R} \to \mathbb{C}$ which satisfy $\psi(\lambda\mu) = \psi(\mu)$ for all $\lambda \in \mathbb{Q} \setminus \{0\}$. When we regard the real numbers as a vector space W over the field of rational numbers this condition is (a slight generalization of) the statement that ψ is a multiplicative function on W. Thus all functions which are also additive, i.e. all *linear* functions $W \to \mathbb{C}$ will give solutions to our constraint. We choose a basis B of W and know from linear algebra that we can obtain a linear function by fixing the values of ψ on the basis arbitrarily (note that the axiom of choice guarantees that there are indeed possible choices for the values of ψ). The basis B has to be uncountable since $\mathbb{R}$ is uncountable.

[7]Even though at first sight this proposal does not seen to bring the desired effect, it has inspired a more radical idea which might bear fruit. We mention it briefly in the outlook.

The space of linear maps $W \to \mathbb{C}$ is therefore non-separable. Since the solution space to our constraint may contain even more elements, it has to be non-separable as well.

Conclusion and Outlook

This thesis started with a (hopefully) pedagogical introduction to the subject of LQG. We began with the derivation of the initial value formulation of general relativity. In chapter 2 we wrote general relativity as a gauge theory. Since it differs substantially from standard field theories on background spacetimes we also presented an alternative quantization program. The first part is devoted to constructing the loop representation of the holonomy-flux Poisson algebra, kinematical geometrical operators and the solution of the Gauß- and diffeomorphism constraints. According to the focus of this thesis we put some emphasis on the kinematical aspects and the methods and tools needed for the calculations in the second part. The Hamiltonian constraint was discussed in sufficient detail so that the constructions in the symmetry-reduced models of part 2 can be compared to those of full LQG. Necessary physical background-material, such as the Lagrangian formulation of general relativity, and mathematical prerequisites are provided as appendices.

The second part focuses on LQC, an application of LQG to cosmological models. After discussing the general framework of the symmetry-reduction we presented two different approaches to the kinematical setting of LQC. We saw that both of them lead to the same dynamical structure. As one of the main purposes of LQC is to learn lessons for full LQG, it is natural that most of the research in the area focuses on dynamical aspects. After all, most open problems in full LQG concern the dynamics. Nevertheless we are convinced that one should also have a solid understanding of the kinematics in order to provide a sound basis for the dynamics.

In chapter 7 we found apparent discrepancies between the two different approaches to the kinematics of LQC. We then presented a proposal for a constraint corresponding to special diffeomorphisms, which we call rescalings, that might lead to a more consistent theory which better mirrors the construction in full LQG. We expected that this constraint would reduce the unreasonably large, non-separable diffeomorphism invariant Hilbert space to a separable space.

Unfortunately our original constraint is too restrictive — it leads to an only two-dimensional diffeomorphism invariant Hilbert space. Even worse we obtain no non-trivial solutions to the Hamiltonian constraint. We then discussed various suggestions on how to modify the original construction. We have to admit that so far none of these suggestions yields the desired result of an infinite dimensional separable Hilbert space. Relaxing the group of rescalings to piecewise constant rescalings did not improve our situation. The corresponding space of solutions is the same as that of the original proposal. On the other

hand, demanding invariance only under rescalings with rational numbers did not reduce the size of the diffeomorphism invariant Hilbert space sufficiently to obtain a countable basis.

We have to conclude that further research is necessary. Let us outline some possible points of departure for future investigations.

- We think that a further refinement of the idea of using piecewise constant rescalings probably has the best chances of leading to the desired result. We could decompose the space of symmetric cylindrical functions into spaces of functions with a fixed number of coefficients, i.e. Cyl_N consists of functions of the form $\sum_{i=1}^{N} \xi_i \exp\left(i\frac{\mu_i c}{2}\right)$. The geometrical picture is a set of N intervals with given length μ_i. The hope is that rescalings might lead to independence of the length of the intervals and a diffeomorphism invariant Hilbert space whose elements are labeled by N, the number of intervals. The Hamiltonian constraint then might act by adding intervals, just as in full LQG it acts by adding edges.
- Another possible approach is to incorporate groupoid structure into the projective description of the Bohr compactification. Whereas in full LQG generalized connections are morphism from the groupoid $\mathcal{P}$ of paths to $SU(2)$, in LQC $\mathcal{P}$ is replaced by the *group* $\mathbb{R}$. The idea is then to replace the length μ of a straight edge by directed intervals (x_1, x_2). Those intervals indeed form a groupoid if one defines $(x_1, x_2) \circ (x_2, x_3) := (x_1, x_3)$ and $(x_1, x_2)^{-1} := (x_2, x_1)$.

Part 3

Appendices

APPENDIX A

Hypersurfaces

When taking into account boundary terms in the action of general relativity, as well as in the 3+1-split of the ADM-formulation, we have to deal with three-dimensional submanifolds Σ of spacetime M. We collect here some important definitions and facts on hypersurfaces that will be needed in the main text. Our exposition follows Hawking and Ellis [25], as well as Wald [42].

1. Definitions

For most of what follows an intuitive understanding of the notion of hypersurface will be sufficient. However, for completeness we give the following technical definition: If S is a p-dimensional manifold and M an n-dimensional manifold with $p < n$, a mapping $\phi\colon S \to M$ is called an immersion, provided that the differential $d\phi(p)$ is an injective linear map at every point of S. If, in addition, the map is a homeomorphism from S onto its image $\phi(S)$, it is called an embedding (the topology on $\phi(S)$ is understood to be the induced topology). The image $\phi(S)$ is then called an embedded submanifold. An embedded submanifold of dimension $n-1$ is a hypersurface. For simplicity, we will suppress the embedding ϕ and just refer to $\phi(S)$ as a hypersurface Σ.

We denote by n^a the unit normal vector to Σ, i.e. $g_{ab}n^a v^b = 0$, $\forall v^b \in T\Sigma$ and $g_{ab}n^a n^b = \pm 1$. Σ will be called timelike if the normal vector is spacelike and vice versa. Note that we explicitly excluded the null case, since most of the following won't work for null hypersurfaces (or at least not that easily).

By the implicit function theorem, Σ is, at least locally, the solution of an equation $f(x) = t$, where t is a constant. Thus, the one-form df satisfies $df(v) = 0$, $\forall v \in T\Sigma$. Since the space of these one-forms is one-dimensional, df has to be proportional to n_a. We thus have the equality

$$n = \alpha df, \tag{A.1}$$

which we will use frequently in the following.

A metric g_{ab} on M will induce a metric h_{ab} on a hypersurface by restriction of the action of g_{ab} to vectors tangential to Σ. If Σ is a spacelike hypersurface of spacetime M, the induced metric will be positive definite and given by the formula

$$h_{ab} = g_{ab} + n_a n_b. \tag{A.2}$$

For timelike Σ the plus sign is replaced by a minus sign and the induced metric will be indefinite. It follows immediately from this formula, that $h^a{}_b$ is a projection operator,

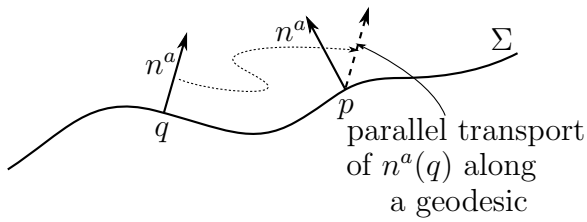

FIGURE 1. The extrinsic curvature describes the failure of n^a at point p to correspond with the parallel transport of the vector n^a at q along a geodesic.

i.e. $h^a{}_b h^b{}_c = h^a{}_c$. Its kernel is the linear span of n^a, and $T\Sigma$ is its image (since obviously $h^a{}_b n^b = 0$, and $h^a{}_b v^b = v^a$ for all tangential v^a). Hence we have the decomposition (for timelike Σ again replace $-$ by $+$)

$$v^a = h^a{}_b v^b - n^a n_b v^b \tag{A.3}$$

$$\omega_a = h_a{}^b \omega_b - n_a n^b \omega_b. \tag{A.4}$$

Similarly, we can project arbitrary tensors into their parts tangential to Σ in all indices. Actually, now it is also obvious that equation (A.2) defines the action of h_{ab} on arbitrary vectors in TM and computes the inner product of their projections on $T\Sigma$.

2. Extrinsic geometry

Intuitively, the variation of the unit normal vector n^a along Σ (i.e. the failure of the normal vector to coincide with the vector obtained by parallel transporting the normal vector from another point) describes the way the hypersurface is embedded in M (see the illustration in figure 1). We thus define the extrinsic curvature by

$$K_{ab} := h_a{}^c \nabla_c n_b. \tag{A.5}$$

Several remarks concerning this definition are in order.

Firstly, note, that we had to project the index c, in order to get an object defined on Σ. But we didn't have to project the second index of $\nabla_c n_b$, because it is already tangential: $n^b \nabla_c n_b = \frac{1}{2}\nabla_c \underbrace{n_b n_b}_{=-1} = 0$.

Secondly, if we are given only one hypersurface, we have to extend the definition of n_a to a neighborhood of Σ, in order to be able to compute the derivatives. However, because of the projection that restricts the derivatives to directions tangential to Σ, K_{ab} will be independent of the special extension used in the definition.

We also note that the extrinsic curvature is symmetric:

$$K_{ab} = h_a{}^c \nabla_c n_b = h_a{}^c h_b{}^d \nabla_c n_d = \alpha h_a{}^c h_b{}^d \nabla_c \nabla_d f = K_{ba}, \tag{A.6}$$

where in the third equality we made use of equation (A.1) and in the last we exploited the fact that ∇_a is torsion-free.

If we are given a foliation of spacetime into hypersurfaces, i.e. $M \cong \mathbb{R} \times \Sigma$, then there is an alternative interpretation of the extrinsic curvature. K_{ab} is the change of the intrinsic geometry of the three-surface, when moving along the integral curves of n^a,

$$K_{ab} = \frac{1}{2} \mathcal{L}_n h_{ab}. \tag{A.7}$$

To prove this formula, we use the well-known identity $i_{[v,w]} = [\mathcal{L}_v, i_w]$ (see e.g. [20]) to conclude that $\mathcal{L}_n h_{ab}$ is tangential,

$$n^a \mathcal{L}_n h_{ab} = \mathcal{L}_n n^a h_{ab} - [n, n]^a h_{ab} = 0.$$

We can then use this fact to write

$$\begin{aligned} \mathcal{L}_n h_{ab} &= h_a{}^c h_b{}^d \mathcal{L}_n h_{cd} = h_a{}^c h_b{}^d \left(\mathcal{L}_n h_{cd} - \mathcal{L}_n (n_c n_d) \right) = \\ &= h_a{}^c h_b{}^d \mathcal{L}_n g_{cd} = h_a{}^c h_b{}^d \left(\nabla_c n_d + \nabla_d n_c \right) = \\ &= K_{ab} + K_{ba} = 2K_{ab}, \end{aligned}$$

where the term that we added in the first line is zero due to $h_a{}^c n_c = 0$ and the Leibniz rule and where we used the symmetry of K_{ab} (equation (A.6)) in the last equality.

We finally remark that formula (A.7) is not really restricted to the case of a congruence of hypersurfaces. This is because for a single surface Σ we can always use Gaussian normal coordinates to define a family of hypersurfaces at least in a neighborhood of Σ. On Σ we chose arbitrary coordinates $x_1, \ldots, x_{n-1}$. For each point p in a neighborhood of Σ there will be a geodesic through p starting at some point q in Σ with tangent vector n^a. The geodesic will be unique if we choose the neighborhood small enough, so that the geodesics do not cross. We label p with the affine parameter t along this geodesic and the $n-1$ coordinates of the point q. Thus, $p = (t, x_1, \ldots x_{n-1})$. A family of surfaces is then given by defining Σ_t by the equation $t =$ constant. Equivalently, $\Sigma_t = \phi_t(\Sigma)$, where ϕ_t is the flow defined by the geodesics. All that remains to check to be able to use this family for formula (A.7) is that the hypersurfaces remain space-like. This is verified by showing that the tangent vector n^a to the geodesics remains orthogonal to the coordinate basis vector-fields $X_1, \ldots X_n$,

$$(n\nabla)\left(n_a X_i^a\right) = n_a (n\nabla) X_i^a = n_a (X_i \nabla) n^a = 0.$$

The first equality follows from the geodesic equation, the second from the fact that the lie-bracket of the coordinate vector-fields n^a and X_i^a vanishes, and the last equality is just the Leibniz rule together with $n_a n^a = -1$.

3. Intrinsic geometry

We already defined in equation (A.2) the induced metric on Σ. Since our hypersurface is embedded in M, it is natural to ask how various metric dependent quantities defined on Σ are related to the corresponding quantities on M.

3.1. Derivative operator. Since Σ is a manifold with metric, there is also a unique derivative operator D_a compatible with h_{ab}.

Lemma A.1.

$$D_c t^{a_1\dots a_k}{}_{b_1\dots b_l} = h^{a_1}{}_{d_1}\dots h_{b_l}{}^{e_l} h_c{}^f \nabla_f t^{d_1\dots d_k}{}_{e_1\dots e_l} \tag{A.8}$$

Proof. First note that similar to the discussion after the definition of K_{ab} in the previous section, we actually have to extend the definition of $t^{a_1\dots a_k}{}_{b_1\dots b_l}$ to a neighborhood of Σ. As before the result will be independent of this extension, because all derivatives are projected to tangential directions.

The axioms of a derivative operator are easily verified. Torsion-freeness is also immediate:

$$\begin{aligned} D_a D_b f &= h_a{}^m h_b{}^n \nabla_m D_n f = h_a{}^m h_b{}^n \nabla_m (h_n{}^p \nabla_p f) \\ &= h_a{}^m h_b{}^n \nabla_m \nabla_n f + h_a{}^m h_b{}^n \nabla_m (n_n n^p) \nabla_p f \\ &= h_a{}^m h_b{}^n \nabla_m \nabla_n f + K_{ab}(n\nabla) f = D_b D_a f, \end{aligned}$$

where we have used equation (A.2) and the compatibility of ∇_a with g_{ab} in the third equality, $h_b{}^n n_n = 0$ in the fourth. The last equality follows from the torsion-freeness of ∇_a and the symmetry of K_{ab} (A.6).

Metric compatibility is verified by

$$D_a h_{bc} = h_a{}^d h_b{}^e h_c{}^f \nabla_d (g_{ef} + n_e n_f) = 0,$$

using the metric compatibility of ∇_a and the Leibniz rule. □

3.2. Riemann tensor. To establish the relation between the Riemann tensors on Σ and M, we use the identity

$$[D_a, D_b]\, v^c = {}^3R^c{}_{dab} v^d \tag{A.9}$$

for arbitrary vectors v^a tangent to Σ. Since we will need the results of this section only for spacelike hypersurfaces, we will restrict ourselves to that case in the following.

We have

$$\begin{aligned} D_a D_b v^c &= h_a{}^m h_b{}^n h^c{}_p \nabla_m (D_n v^p) = h_a{}^m h_b{}^n h^c{}_p \nabla_m (h_n{}^q h^p{}_r \nabla_q v^r) \\ &= h_a{}^m h_b{}^q h^c{}_r \nabla_m \nabla_q v^r + h_a{}^m h_b{}^n h^c{}_r \nabla_m h_n{}^q \nabla_q v^r + h_a{}^m h_b{}^q h^c{}_p \nabla_m h^p{}_r \nabla_q v^r. \end{aligned}$$

The first term already yields the Riemann tensor on M after we antisymmetrize, whereas we will simplify the second term using the definition of the induced metric (A.2), together

with metric compatibility of ∇_a, $h_b{}^n n_n = 0$ and the definition of K_{ab} (A.5):

$$h_a{}^m h_b{}^n h^c{}_r \nabla_m h_n{}^q \nabla_q v^r = h_a{}^m h_b{}^n h^c{}_r \nabla_m (n_n n^q) \nabla_q v^r = K_{ab} h^c{}_r n^q \nabla_q v^r .$$

Proceeding in a similar way and taking into account that v^a is tangential (i.e. $n_r v^r = 0$), we obtain for the third term:

$$h_a{}^m h_b{}^q h^c{}_p \nabla_m h^p{}_r \nabla_q v^r = K_a{}^c h_b{}^q n_r \nabla_q v^r = -K_a{}^c h_b{}^q v^r \nabla_q n_r = -K_a{}^c K_{br} v^r .$$

Collecting terms, antisymmetrizing and inserting a projector in the first term ($v^s = h^s{}_d v^d$ due to v^s being tangential), we arrive at

$${}^3R^c{}_{dab} v^d = h_a{}^m h_b{}^q h^c{}_r h^s{}_d R^r{}_{smq} v^d - K_a{}^c K_{bd} v^d + K_b{}^c K_{ad} v^d$$

where we have arranged the indices to isolate v^d. Since this equation has to be valid for all tangential v^d, we arrive at Gauß's equation

$${}^3R^c{}_{dab} = h^c{}_r h_d{}^s h_a{}^m h_b{}^q R^r{}_{smq} - K_a{}^c K_{bd} + K_b{}^c K_{ad} . \tag{A.10}$$

In view of the 3+1-split we are going to perform in chapter 1, we will make two contractions to obtain the relation between the curvature scalars. By using equation (A.2) for $h_r{}^m$, we obtain

$$\begin{aligned} {}^3R_{ab} &= h_r{}^m h_a{}^s h_b{}^q R^r{}_{smq} - K K_{ba} + K_b{}^c K_{ca} \\ &= h_a{}^s h_b{}^q R_{sq} + n_r n^m h_a{}^s h_b{}^q R^r{}_{smq} - K K_{ba} + K_b{}^c K_{ca} \end{aligned}$$

for the Ricci tensor. Another contraction with h^{ab} yields the curvature scalar,

$$\begin{aligned} {}^3R &= h^{sq} R_{sq} + n_r n^m h^{sq} R^r{}_{smq} - K^2 + K^{ac} K_{ca} \\ &= R + 2 n^m n^r R_{mr} + n^r n^m n^s n^q R_{rsmq} - K^2 + K_{ab} K^{ab} \\ &= R + 2 n^m n^r R_{mr} - K^2 + K_{ab} K^{ab} , \end{aligned}$$

again by using the definition of the induced metric, equation (A.2) twice and taking into account the antisymmetry properties of the Riemann tensor in the last equality.

We can take care of the disturbing second term by inserting the definition of the Riemann tensor and using the Leibniz rule,

$$\begin{aligned} n^m n^r R_{mr} &= n^m n^r R^c{}_{mcr} = n^r \left[\nabla_c , \nabla_r \right] n^c = n^r \nabla_c \nabla_r n^c - n^r \nabla_r \nabla_c n^c \\ &= -\nabla_c n^r \nabla_r n^c + \nabla_r n^r \nabla_c n^c + \nabla_c \left(n^r \nabla_r n^c \right) - \nabla_r \left(n^r \nabla_c n^c \right) . \end{aligned}$$

With the identities

$$\nabla_a n^a = g^{ab} \nabla_a n_b = h^{ab} \nabla_a n_b - n^a n^b \nabla_a n_b = K$$

and

$$\begin{aligned} \nabla_c n^r \nabla_r n^c &= g^{ar} g^{bc} \nabla_c n_a \nabla_r n_b = \left(h^{ar} - n^a n^r \right) \left(h^{bc} - n^b n^c \right) \nabla_c n_a \nabla_r n_b \\ &= \left(h^{ar} - n^a n^r \right) h^{bc} \nabla_c n_a \nabla_r n_b = h^{ar} h^{bc} \nabla_c n_a \nabla_r n_b \\ &= K^b{}_a K^a{}_b = K_{ab} K^{ab} , \end{aligned}$$

we can further simplify this expression:

$$n^m n^r R_{mr} = -K_{ab}K^{ab} + K^2 + \nabla_c\left((n\nabla)n^c\right) - \nabla_r\left(n^r K\right).$$

Our final result for the curvature scalar is[1]

$$^3R = R + K^2 - K_{ab}K^{ab} + 2\nabla_a\left((n\nabla)n^a\right) - 2\nabla_a\left(n^a K\right). \tag{A.11}$$

3.3. Volume element. To even properly define a volume element, the underlying manifold should be orientable. Wald [42] gives some basic definitions, but for a thorough understanding, Warner's classic book [43] is highly recommended.

We define a manifold of dimension n to be orientable provided that there exists a continuous nowhere vanishing n-form ϵ. Those forms are divided into two equivalence classes by the relation $\epsilon \sim \epsilon'$ iff $\epsilon = f\epsilon'$, with a strictly positive function f. An orientation is then a choice of one of the two equivalence classes. We call the forms in this class right-handed or oriented.

Let $\{E_1, \dots, E_n\}$ be a basis of T_pM and $\{e^1, \dots, e^n\}$ its dual basis. We call this basis oriented if $e^1 \wedge \cdots \wedge e^n$ is oriented. A coordinate-system $(x^1, \dots, x^n)$ is oriented, if $dx_1 \wedge \cdots \wedge dx_n$ is. Since the reader might might be familiar with the definition of orientability via the existence of an oriented atlas (i.e. an atlas where all the coordinate changes have positive Jacobi determinant), we mention that this definition is equivalent to ours. Warner [43] gives a proof in Proposition 4.2.

A hypersurface Σ of an oriented manifold M need not be orientable (consider e.g. the Möbius-strip embedded in $\mathbb{R}^3$), and even if it is, there is no natural notion on how M induces an orientation on Σ [2]. However, it is always possible to define an orientation of the boundary of a manifold, once we have agreed on one of the two possible ways to do so. A manifold with boundary is locally isomorphic to a subset of $\mathbb{R}^n_- := \{(x^1, \dots, x^n) \in \mathbb{R}^n \mid x^1 \leq 0\}$, just as a manifold without boundary is locally isomorphic to a subset of $\mathbb{R}^n$. The boundary ∂M will be given locally by the equation $x^1 = 0$ and ∂_1 will be called outward-pointing. $(x^2, \dots, x^n)$ will then be a local coordinate system on ∂M and we agree to call it right-handed provided that $(x^1, \dots, x^n)$ is right handed. The more customary definition is probably by defining a basis $(E_2, \dots, E_n)$ of $T_p\partial M$ to be oriented iff for every outward-pointing vector v, $(v, E_2, \dots, E_n)$ is an oriented base of T_pM. Since ∂_1 is outward pointing, these definitions are easily seen to be equivalent.

On an oriented manifold, we can define the integral of an n-form α. If M is covered by a single right-handed coordinate system $(x^1, \dots, x^n)$ and $\alpha = a(x^1, \dots, x^n)dx^1 \wedge \cdots \wedge dx^n$, we define its integral by

$$\int \alpha := \int a(x_1, \dots, x_n)dx^1 \dots dx^n. \tag{A.12}$$

[1]For timelike Σ all terms except the four-dimensional curvature scalar R receive an extra sign.

[2]An embedded submanifold can be oriented by requiring that the differential of the embedding carries oriented bases into oriented bases. Nevertheless, given the hypersurface as a subset of M, there is in general more than one way to define the corresponding embedding map!

For a left-handed coordinate-system we define the integral with an overall minus-sign. Note that the Jacobi determinant from the transformation behavior of α just gives the usual transformation law of integrals over $\mathbb{R}^n$, so this definition is coordinate-independent. Also observe that orientability plays a crucial role here, since the transformation law requires the determinant to be positive! For our purposes this definition of the integral will be sufficient, although we skipped over many subtleties (most notably the fact that usually M is *not* covered by a single coordinate-system). For those we refer to the extensive literature on differentiable manifolds, most notably [43]. Also most books on general relativity treat this topic at least in some detail.

If we are given a volume element on an orientable manifold, i.e. a continuous, nowhere vanishing n-form ω, we can use it to integrate functions, simply by defining

$$\int f := \int f\omega.$$

Note that a volume element automatically selects an orientation by requiring that it be oriented.

Given a metric and orientation, a natural volume element is defined in every right-handed coordinate-system by

$$\omega_g := \sqrt{|\det g|}dx^1 \wedge \cdots \wedge dx^n = e^1 \wedge \cdots \wedge e^n, \tag{A.13}$$

where $e^1, \ldots, e^n$ is the dual basis to an oriented orthonormal basis $E_1^a, \ldots E_n^a$. This definition is easily seen to be independent of the special oriented coordinate-system in use, but it picks up an additional minus sign for left-handed coordinate-systems. Thus, the metric can not be used to define an orientation!

We will now come to the issue of how the natural volume elements on M and Σ are related. Because of the above mentioned impossibility of inducing an orientation on Σ, we have to agree on its orientation first. We fix the sign of the normal vector n^a and let a basis $(E_2^a, \ldots, E_n^a)$ of $T_p\Sigma$ be oriented iff $(n^a, E_2^a, \ldots, E_n^a)$ is oriented. Thus, in the case where Σ is the boundary ∂M of M, we just have to fix the sign of n^a such that it be outward-pointing in order to be consistent with the previous definition.[3] We then have the following lemma.

Lemma A.2. *The volume elements ω_g on M and ω_h on Σ are related by*

$$\omega_g = \sigma n \wedge \omega_h, \tag{A.14}$$

where σ is the sign of $n_a n^a$.

[3]There are two more special cases in which a natural choice for the sign of n^a and therefore for the orientation of Σ is at hand. In the Euclidean case, for a hypersurface that divides the manifold into a compact and a non-compact part, it is customary to call the non-compact part "outside" and choose n^a to be outward-pointing.
In relativity, when we are given a time-orientation, one usually takes the normal vectors to a spacelike hypersurface to be future-directed.

Proof. The proof is most easily done using an adapted oriented orthonormal basis $(E_1^a, \ldots, E_n^a)$ with dual basis $(e^1, \ldots, e^n)$ and $E_1^a := n^a$. This implies $n_a = \sigma e_a^1$, since $e_a^1 E_j^a = \delta_j^1$ but $n_a E_j^a = \sigma \delta_j^1$. We can thus easily read off the desired relation from (A.13). □

4. Stokes's theorem

Finally we cast Stokes's theorem into a form that is often used in general relativity. We will use the method employed by Carroll [19].

Theorem A.3 (Stokes). *For an n-dimensional compact oriented manifold M with boundary ∂M, and a differentiable $(n-1)$-form α on M,*

$$\int_M d\alpha = \int_{\partial M} \alpha. \tag{A.15}$$

For a proof see e.g. [43]. A convenient way to rewrite this equation is by noticing that any $(n-1)$-form α can be written as the Hodge-dual of a one-form v,

$$\alpha_{a_1 \ldots a_{n-1}} = (*v)_{a_1 \ldots a_{n-1}} = \epsilon^b{}_{a_1 \ldots a_{n-1}} v_b = \epsilon_{b a_1 a_{n-1}} v^b.$$

Using metric compatibility of the Levi-Civita tensor, the integrand of the left-hand side of Stokes's theorem can then be written as

$$(d\alpha)_{c a_1 \ldots a_{n-1}} = n \nabla_{[c} \left(\epsilon_{|b| a_1 \ldots a_{n-1}]} v^b \right) = n \epsilon_{b[a_1 a_{n-1}} \nabla_{c]} v^b.$$

Since any n-form is proportional to ϵ, we have

$$d\alpha = f\epsilon = *f. \tag{A.16}$$

The function f can be found by taking the Hodge-dual of this equation, giving (remember equation (N.10))

$$\begin{aligned} f &= (-1)^s * *f = (-1)^s * d\alpha = (-1)^s \frac{1}{n!} \epsilon^{c a_1 \ldots a_{n-1}} \left(n \epsilon_{b[a_1 \ldots a_{n-1}} \nabla_{c]} v^b \right) \\ &= (-1)^s \frac{1}{(n-1)!} (n-1)! \delta_b^c (-1)^s \nabla_c v^b = \nabla_b v^b. \end{aligned}$$

Inserting into equation (A.16), and remembering that the Levi-Civita tensor equals the volume element, we obtain

$$d\alpha = \nabla_b v^b \omega_g.$$

The integrand of the right-hand side is easily found by taking the inner product of v^a with equation (A.14),

$$\alpha = v \lrcorner \omega_g = \sigma \left(v \lrcorner n \right) \omega_h - \sigma n \wedge \left(v \lrcorner \omega_h \right).$$

The second term on the right vanishes when acting on vectors tangent to ∂M. We therefore have

$$\alpha = \sigma n_a v^a \omega_h.$$

Putting our results together, we can write Stokes's theorem in the very useful form

$$\int_M \nabla_a v^a \omega_g = \sigma \int_{\partial M} n_a v^a \omega_h, \tag{A.17}$$

which is also known as Gauß's law. Remember that, as said before lemma A.2, n^a needs to be outward-pointing in order to get the correct orientation on ∂M.

APPENDIX B

Lagrangian Formulation of General Relativity

Since most attempts at quantizing a physical theory start from an action principle (or from a Hamiltonian formulation, which in turn is most conveniently derived from the action), we will derive here the Einstein-Hilbert-Hawking action for general relativity starting from the field equations in vacuum. A good reference on this topic is appendix E of Wald's book [42], where one can also find directions on how to include matter-fields.

Further motivation for constructing Lagrangian and Hamiltonian formulations of general relativity is provided by the fact that the former allows for a very natural definition of the stress-energy tensor via Noether's theorem, while the latter provides considerable insight into the dynamics (see chapter 1).

1. Preliminaries

Our goal is thus to find a real-valued functional $S[g_{ab}]$ of the metric, such that its minimum is attained at that value of g_{ab} which satisfies the field equations $G_{ab} = 0$. To find the minimum, we have to set the derivative of S to zero. Since S is a map between the spaces $\mathcal{M}$ of all metrics on M and $\mathbb{R}$, this is generally not an easy task. However, we can simplify things by considering instead all one-parameter curves $g_{ab}(\lambda)$ in $\mathcal{M}$ starting at g_{ab}, that is $g_{ab}(0) = g_{ab}$. For a minimum at g_{ab}, the derivative of $S[g_{ab}(\lambda)]$ at $\lambda = 0$ will then vanish. Denoting $\delta := \frac{d}{d\lambda}\big|_{\lambda=0}$ and dropping the dependence on λ, this condition can be written as $\delta S[g_{ab}] = 0$.

To summarize, we have to find a function S, such that

$$\delta S[g_{ab}] = -\int G^{ab}\delta g_{ab}\omega_g. \tag{B.1}$$

We will start by providing some formulas, that will come in handy later on.

Lemma B.1.

$$\delta\omega_g = \frac{1}{2}g^{ab}\delta g_{ab}\omega_g \tag{B.2}$$

Proof. Since locally we have $\omega_g = \sqrt{|\det g_{\mu\nu}|}d^4x$, we can prove the lemma by just working with matrices for the coordinate-components of g_{ab}.

For an arbitrary matrix M, the determinant is given by $\det M = \det(\Delta) = \prod_{i=1}^n \lambda_i$, where $\Delta = UMU^{-1}$ is diagonal with the eigenvalues λ_i as its entries. Thus, by the Leibniz

rule and because the trace is independent of the basis chosen to evaluate it,

$$\begin{aligned}\delta \det M &= \sum_{i=1}^{n} \lambda_1 \dots \delta\lambda_i \dots \lambda_n = \sum_{i=1}^{n} \frac{\delta\lambda_i}{\lambda_i} \det M \\ &= \operatorname{tr}\left(\Delta^{-1}\delta\Delta\right) \det M, = \operatorname{tr}\left(M^{-1}\delta M\right) \det M,\end{aligned}$$

which leads to

$$\delta\sqrt{\det M} = \frac{1}{2\sqrt{\det M}} \operatorname{tr}(M^{-1}\delta M) \det M = \frac{1}{2} \operatorname{tr}\left(M^{-1}\delta M\right) \sqrt{\det M}. \qquad \square$$

To shorten notation, we will define $\delta g^a{}_b := g^{am}\delta g_{mb}$ and similarly when we raise two indices. We will just have to keep in mind that S is a functional of g_{ab}, hence the variation by default always concerns the metric. In the rare cases when we really want to compute the variation of the inverse metric, we will explicitly use brackets, as in the following lemma. We will use this notation later on also for other tensors. When we write a variation of a tensor, whose indices are *not* in their natural position, we will really mean the variation of the tensor with the indices in their *natural* position, with indices being lowered or raised after the variation. An example should suffice to get the idea: $\delta C^a{}_b{}^c := g^{cm}\delta C^a{}_{bm}$. Note that in combination with our convention for contractions, this also defines $\delta g := g^{ab}\delta g_{ab}$ and similar for other tensors.

Lemma B.2.

$$\delta(g^{ab}) = -g^{am}g^{bn}\delta g_{mn}. \tag{B.3}$$

Proof. Varying the identity $g^{ab}g_{bc} = \delta^a_c$ gives

$$\delta(g^{ab})g_{bc} + g^{ab}\delta g_{bc} = 0.$$

Raising the index c with g^{cd} leads to

$$\delta(g^{ab})\delta^d_b = -g^{ab}g^{cd}\delta g_{bc},$$

which is the desired formula (after a renaming of the indices). $\square$

Note that the inverse metric depends nonlinearly on the metric via its first minors, so that the variation of g^{ab} is a highly nonlinear function of δg_{ab}.

Since ∇_a depends on the metric via the compatibility condition $\nabla_a g_{bc} = 0$, we can also ask for its variation. If we use the natural pointwise definition $(\delta\nabla)_c v^a := \delta(\nabla_c v^a)$, then $\delta\nabla$ will be a new derivative operator.

Lemma B.3. *For vectors v^a not depending on λ the variation of the connection is given by*

$$(\delta\nabla)_c v^a = \delta C^a{}_{bc} v^b, \tag{B.4}$$

where the variation of the difference tensor is

$$\delta C^a{}_{bc} = \frac{1}{2}\left[\nabla_c \delta g^a{}_b + \nabla_b \delta g^a{}_c - \nabla^a \delta g_{bc}\right] \tag{B.5}$$

Proof. Let ${}^\lambda\nabla$ be compatible with $g_{ab}(\lambda)$ and ∇ with g_{ab} (for this proof, quantities that are not labeled with λ will mean the value at $\lambda = 0$). We can express ${}^\lambda\nabla$ in terms of ∇ and a difference tensor $C^a{}_{bc}(\lambda)$,

$$ {}^\lambda\nabla_c v^a = \nabla_c v^a + C^a{}_{bc}(\lambda) v^b. \tag{B.6} $$

Variation of this equation gives the result (B.4). The variation of the difference tensor is also easily computed via

$$ C^a{}_{bc}(\lambda) = \frac{1}{2} g^{am}(\lambda) \left[\nabla_c g_{mb}(\lambda) + \nabla_b g_{cm}(\lambda) - \nabla_m g_{bc}(\lambda) \right], $$

by observing that every time the variation acts on the inverse metric, after evaluating at $\lambda = 0$ we are left with $\nabla_a g_{bc}$, which is zero by metric compatibility. □

The action of $\delta\nabla$ on tensors with arbitrary indices is computed in the usual way. We also have a Leibniz rule for the action on tensor fields depending on λ:

$$ (\delta\nabla)_c n^a = (\delta\nabla)_c n^a + \nabla_c \delta n^a, \tag{B.7} $$

as can easily be seen by using equation (B.6) with n^a in place of v^a and taking into account $C^a{}_{bc}(0) = 0$. Finally, note that $\delta\nabla$ is *not* the derivative operator compatible with δg_{ab} (which is not even a metric!).

2. Hilbert-term

After those preparations, it is now time to derive the action of general relativity. We will start with equation (B.1) and try to write it as the variation of a function that we can then identify as $S[g_{ab}]$.

$$ \begin{aligned} \int \left(R^{ab} - \frac{1}{2} R g^{ab} \right) \delta g_{ab} \omega_g &= - \int R_{ab} \delta(g^{ab}) \omega_g - \int R \delta\omega_g \\ &= \int \left(-\delta R + g^{ab} \delta R_{ab} \right) \omega_g - \int R \delta\omega_g \\ &= \int \left(-\delta \left(R\omega_g \right) + R\delta\omega_g + g^{ab} \delta R_{ab} \omega_g - R\delta\omega_g \right) \\ &= \delta \left(- \int R\omega_g \right) + \int g^{ab} \delta R_{ab} \omega_g. \end{aligned} \tag{B.8} $$

We used Lemmas B.2 and B.1 in the first line and made frequent use of the Leibniz rule from there on.

The first term on the right-hand side already has the desired form, whereas the second term has to be further investigated. To that end, we observe that we can write $\delta R_{ab} = \delta^d_c \delta R^c{}_{adb}$ and use the well-known formula for the Riemann-tensor,

$$ R^c{}_{adb}(\lambda) = \nabla_d C^c{}_{ab}(\lambda) - \nabla_b C^c{}_{ad}(\lambda) + C^c{}_{md}(\lambda) C^m{}_{ab}(\lambda) - C^c{}_{mb}(\lambda) C^m{}_{ad}(\lambda). $$

Varying this expression, we see that the terms quadratic in the difference-tensor vanish due to the Leibniz rule and $C^a{}_{bc}(0) = 0$,

$$\delta R^c{}_{adb} = \nabla_d \delta C^c{}_{ab} - \nabla_b \delta C^c{}_{ad}.$$

We can now use equation (B.5) to simplify the second term on the right hand side in (B.8). We will use the notation introduced just before Lemma B.2 (but we have to be careful not to denote $g^{ab}\delta R_{ab}$ by δR, since the latter already means variation of the curvature scalar).

$$\begin{aligned} g^{ab}\delta R_{ab} &= g^{ab}\delta R^c{}_{acb} = g^{ab}\left(\nabla_c \delta C^c{}_{ab} - \nabla_b \delta C^c{}_{ac}\right) = \nabla_c \delta C^c{}_a{}^a - \nabla_b C^{cb}{}_c \\ &= \frac{1}{2}\nabla_c\left(\nabla^a \delta g^c{}_a + \nabla_a \delta g^{ca} - \nabla^c \delta g_a{}^a\right) - \frac{1}{2}\nabla_b\left(\nabla_c \delta g^{cb} + \nabla^b \delta g^c{}_c - \nabla^c \delta g^b{}_c\right) \\ &= \frac{1}{2}\nabla_c\left(2\nabla_a \delta g^{ac} - \nabla^c \delta g\right) - \frac{1}{2}\nabla_b \nabla^b \delta g = \nabla_a \nabla_b \delta g^{ab} - \nabla^2 \delta g \\ &= \nabla_a\left(\nabla_b \delta g^{ab} - \nabla^a \delta g\right). \end{aligned} \tag{B.9}$$

We thus arrive at

$$\int G^{ab}\delta g_{ab}\omega_g = -\delta\left(\int R\omega_g\right) + \int \nabla_a\left(\nabla_b \delta g^{ab} - \nabla^a \delta g\right)\omega_g.$$

Using the Gauß's law form of Stokes's theorem (A.17), we can rewrite the second term as a surface integral to obtain

$$\int_M G^{ab}\delta g_{ab}\omega_g = -\delta\left(\int_M R\omega_g\right) + \sigma\int_{\partial M} n_a\left(\nabla_b \delta g^{ab} - \nabla^a \delta g\right)\omega_h, \tag{B.10}$$

with the usual convention that n^a is outward-pointing.

The fact that one usually fixes all the fields on the boundary when performing variations caused Einstein and Hilbert to drop the boundary term and consider the action

$$S = \int R\omega_g. \tag{B.11}$$

Note however, that in order for the boundary term to vanish, we not only have to keep the metric constant on ∂M, but also demand that the variation of its derivative vanish. This is of course not desirable, hence we will try to modify the action by adding a term (living only on ∂M) to compensate the annoying boundary term.

3. Hawking boundary term

As already remarked, we will now demand $\delta g_{ab}|_{\partial M} = 0$ but make no assumptions on $(n\nabla)\,\delta g_{ab}|_{\partial M}$. Note that this still implies that derivatives of δg_{ab} in directions tangential to ∂M vanish.

We will start by rewriting the integrand of the boundary term,

$$\begin{aligned} n_a\left(\nabla_b \delta g^{ab} - \nabla^a \delta g\right) &= n^a g^{bc}\left(\nabla_c \delta g_{ab} - \nabla_a \delta g_{bc}\right) \\ &= n^a h^{bc}\left(\nabla_c \delta g_{ab} - \nabla_a \delta g_{bc}\right) = -h^{bc}(n\nabla)\delta g_{bc}. \end{aligned}$$

We used equation (A.2) in the second equality to replace g^{bc} by h^{bc}, where the two terms with triple products of the normal vector cancel. For the last equality, we used the vanishing of the tangential derivatives of δg_{ab}.

The way to rewrite this as a variation is to consider the quantity $K := h^{ab}K_{ab}$.

As h_{ab} is just g_{ab} restricted to ∂M, the variation $\delta h_{ab}|_{\partial M}$ vanishes as well and by Lemma B.3 we have on the boundary of M

$$\begin{aligned}\delta(K) &= h^{ab}\delta K_{ab} = h^{ab}h_a{}^c(\delta\nabla)_c n_b + h^{ab}h_a{}^c\delta n_b = -h^{bc}\delta C^a{}_{bc}n_a \\ &= -h^{bc}n_a\frac{1}{2}\left(\nabla_c\delta g^a{}_b + \nabla_b\delta g^a{}_c - \nabla^a\delta g_{bc}\right) = \frac{1}{2}h^{bc}(n\nabla)\delta g_{bc}.\end{aligned}$$

We could cancel the δn_b-term because in $n_b = \alpha\nabla_b t$ we can use a fixed derivative operator ∇_b (it acts only on a function). Thus $\delta n_b = \frac{\delta\alpha}{\alpha}n_b$ and therefore $h^{ab}\delta n_b = \frac{\delta\alpha}{\alpha}h^{ab}n_b = 0$. In the last equality we used the vanishing of tangential derivatives of δg_{ab}.

Combining the last two equations and taking into account $\delta\omega_h = 0$ we finally arrive at the following expression for the boundary term:

$$\int_{\partial M} n_a\left(\nabla_b\delta g^{ab} - \nabla^a\delta g\right) = -2\int_{\partial M}\delta K\omega_h = \delta\left(-2\int_{\partial M}K\omega_h\right).$$

Putting our results together, we have

$$\delta\left[\int_M R\omega_g + 2\sigma\int_{\partial M}K\omega_h\right] = -\int_M G_{ab}\delta g^{ab}\omega_g,$$

with the convention that the normal vector (used in the definition of K) is chosen to be outward-pointing. The expression in square brackets is known as Einstein-Hilbert-Hawking action,

$$S_{\text{EHH}} = \int_M R\omega_g + 2\sigma\int_{\partial M}K\omega_h. \tag{B.12}$$

APPENDIX C

Symplectic geometry

For the convenience of the reader we collect here some basic facts about symplectic manifolds that will be needed in the main text. An exhaustive source on this topic is [1].

A *symplectic manifold* is defined as a pair (M, ω), where M is a differentiable manifold and ω is a closed ($d\omega = 0$) and non-degenerate 2-form on M. The condition of non-degeneracy can be expressed equivalently by

- If $\omega_p(X_p, Y_p) = 0$ for all tangent vectors X_p at p, then also $Y_p = 0$,
- The coordinate matrix ω_{ij} satisfies $\det \omega_{ij} \neq 0$,
- $X_p \mapsto X_p \lrcorner \omega$ is an isomorphism at every point $p \in M$.

It immediately follows that the dimension m of M has to be even:

$$0 \neq \det \omega_{ij} = \det(-\omega_{ji}) = (-1)^m \det \omega_{ji} = (-1)^m \det \omega_{ij},$$

hence $1 = (-1)^m$, which can only hold for even m. We will therefore denote the dimension of M by $2n$.

We also define the Poisson tensor P^{ab} as the inverse of the symplectic form,

$$P^{ab}\omega_{ac} := \delta^b_c. \tag{C.1}$$

Note that the antisymmetry of the symplectic form implies that the Poisson tensor is antisymmetric as well.

Vector fields X satisfying $\mathcal{L}_X\omega = 0$ are called *locally Hamiltonian*, whereas *Hamiltonian* vector fields are such that $X \lrcorner \omega$ is exact ($X \lrcorner \omega = -df$ for some function f). Because of $\mathcal{L}_X\omega = d(X \lrcorner \omega) + X \lrcorner d\omega$ and $d\omega = 0$, every *Hamiltonian* vector field is indeed *locally Hamiltonian*. Conversely, the local version of the Poincare Lemma tells us that every *locally Hamiltonian* vector field locally is a *Hamiltonian* vector field.[1]

The condition for Hamiltonian vector fields can also be rewritten in terms of the Poisson-tensor,

$$X^a_f = -P^{ab}\partial_b f. \tag{C.2}$$

From this formulation, we see that for every differentiable function f we get a Hamiltonian vector-field X_f. This situation should be compared to the usually very scarce solutions of the Killing-equation $\mathcal{L}_X g_{ab} = 0$ for Riemannian manifolds.

[1]The sign in $X \lrcorner f = -df$ is of course conventional and a lot of authors use $+$ instead of $-$ here. Also the definition of the Poisson bracket (C.3) might acquire an extra sign depending on the taste of the author. We choose our conventions in order to arrive at (C.7) and (C.9).

We can use the existence of Hamiltonian vector fields to define Poisson-brackets between functions f and g on M,

$$\{f,g\} := \omega\left(X_g, X_f\right) = df\left(X_g\right) = X_g^a \partial_a f = P^{ab}\partial_b f \partial_a g. \tag{C.3}$$

The locally Hamiltonian vector fields form a sub Lie algebra of all vector fields with the commutator as multiplication, $(\mathcal{X}(M), [\cdot,\cdot])$. The proof is immediate, for if X and Y are locally Hamiltonian, then also $\mathcal{L}_{[X,Y]}\omega = [\mathcal{L}_X, \mathcal{L}_Y]\,\omega = 0$.

It is also easy to see that in this sub Lie algebra the Hamiltonian vector-fields form an ideal. If X is Hamiltonian (thus in particular locally Hamiltonian with $\mathcal{L}_X\omega = 0$) and Y is locally Hamiltonian, then

$$\begin{aligned}
[X,Y] \lrcorner\, \omega &= (\mathcal{L}_X Y) \lrcorner\, \omega = \mathcal{L}_X\left(Y \lrcorner\, \omega\right) - Y \lrcorner\, \mathcal{L}_X\omega \\
&= X \lrcorner\, d\left(Y \lrcorner\, \omega\right) + d\left(X \lrcorner\, Y \lrcorner\, \omega\right) \\
&= X \lrcorner\, \underbrace{\left(d\left(Y \lrcorner\, \omega\right) + Y \lrcorner\, d\omega\right)}_{=\mathcal{L}_Y\omega=0} + d\left(X \lrcorner\, Y \lrcorner\, \omega\right),
\end{aligned}$$

by using the product rule for the Lie derivative in the second equality and adding a term $X \lrcorner\, Y \lrcorner\, d\omega = 0$ in the fourth. Thus we have the multiplication property of the ideal,

$$[X,Y] = X_{-X \lrcorner\, Y \lrcorner\, \omega}.$$

Note that this equation also shows that the mapping $X\colon f \mapsto X_f$ is a Lie algebra (anti-)isomorphism from $(C^\infty(M), \{\cdot,\cdot\})$ to $(\mathcal{X}(M), [\cdot,\cdot])$,

$$[X_f, X_g] = X_{\{g,f\}}. \tag{C.4}$$

Since the Lie derivative of a scalar coincides with the directional derivative, we have

$$\{f,g\} = X_g(f) = \mathcal{L}_{X_g} f.$$

Using this equation in

$$\mathcal{L}_{X_{\{f,g\}}} h = \mathcal{L}_{[X_g, X_f]} h = [\mathcal{L}_{X_g}, \mathcal{L}_{X_f}] h$$

we arrive at the Jacobi identity

$$\{\{f,g\},h\} = \{\{f,h\},g\} + \{f,\{g,h\}\}. \tag{C.5}$$

Note that this identity can be seen as a Leibniz rule for the derivation $\{\cdot,h\}$ on the Lie Algebra $(C^\infty(M), \{\cdot,\cdot\})$.

A particularly interesting case of a symplectic manifold, especially for physics, is the cotangent bundle T^*M of a given manifold M. For the benefit of the reader who is not well skilled in differential geometry we give a crude definition. If q^i are coordinates in a region U of the n-dimensional manifold M, then $U \times \mathbb{R}^n$ defines a coordinate neighborhood of the cotangent bundle with coordinates (q^i, p_i). The cotangent bundle therefore has twice the dimension of M. In the overlap of two regions U and $\widetilde{U}$, the coordinates transform according to $(\widetilde{q}^i, \widetilde{p}_i) = (\widetilde{q}^i(q^j), \frac{\partial q^j}{\partial \widetilde{q}^i} p_j)$. Observe that while the coordinates q^i in the base manifold transform in a nonlinear way, the fiber coordinates p_i transform linearly. To

see that this definition makes sense, one only has to check consistency in a triple overlap region, $\widetilde{\widetilde{p}}_i = \frac{\partial q^j}{\partial \widetilde{\widetilde{q}}^i} p_j = \frac{\partial q^j}{\partial \widetilde{q}^k} \frac{\partial \widetilde{q}^k}{\partial \widetilde{\widetilde{q}}^i} p_j = \frac{\partial \widetilde{q}^k}{\partial \widetilde{\widetilde{q}}^i} \widetilde{p}_k$. Consistency for the coordinates q^i in the base manifold follows from the fact that M already is a manifold.

The fact that makes the cotangent bundle interesting is that it carries a natural symplectic structure.[2] In one of the above coordinate systems we have the one-form

$$\theta := p_i dq^i.$$

Actually this already defines a natural one-form globally, because the definition is independent of the chosen coordinate-system,

$$\widetilde{\theta} = \widetilde{p}_i d\widetilde{q}^i = \frac{\partial q^k}{\partial \widetilde{q}^i} p_k \frac{\partial \widetilde{q}^i}{\partial q^j} dq^j = \theta.$$

This one-form then defines the symplectic form

$$\omega := d\theta = dp_i \wedge dq^i.$$

ω is obviously closed since $d^2 = 0$ and non-degenerate — in the above coordinates it is given by[3]

$$\omega_{mn} = \begin{pmatrix} \mathbb{0} & -\mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix}. \tag{C.6}$$

We remark that this construction is even universal in the sense that every symplectic manifold locally is of this structure. This is the content of the theorem of Darboux, which is proved for example in [40]. The above coordinates are therefore called Darboux coordinates.

The universality of Darboux coordinates allows us to calculate Poisson brackets easily. By inverting the symplectic form (C.6) we obtain the Poisson tensor (mind that the Poisson tensor actually is *minus* the inverse of the symplectic form, so it is again given by the matrix (C.6)) . Inserting into the last expression of (C.3) then yields the familiar formula for calculating Poisson brackets in Darboux coordinates

$$\{f, g\} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}. \tag{C.7}$$

In the main text we will be concerned with field theory, so we also give the generalization to infinitely many degrees of freedom. As for most formulas, we just have to view the field variable as a (continuous) degree of freedom and replace the sum by an integral,

$$\{f, g\} = \int \left(\frac{\delta f}{\delta q_{ab}} \frac{\delta g}{\delta \widetilde{\pi}^{ab}} - \frac{\delta f}{\delta \widetilde{\pi}^{ab}} \frac{\delta g}{\delta q_{ab}} \right). \tag{C.8}$$

[2]Actually, θ can be defined in a coordinate-free manner by its action on arbitrary vectors v tangent to T^*M. v is projected to a vector $\pi' v$ tangent to M by the differential of the canonical projection $\pi\colon T^*M \to M$. If v is tangent to T^*M at the point p, we obtain a real number by acting with p on $\pi' v$. Thus $\theta := p(\pi' v)$.

[3]As usual, the first index m denotes the row and the second index n denotes the column.

Symplectic manifolds are the natural home of Hamiltonian mechanics. The notation for Darboux coordinates was already chosen in order to anticipate this. A Hamilton function H is simply a function on the symplectic manifold M, and the integral curves of the corresponding Hamiltonian vector field X_H are called dynamical trajectories. Employing Darboux coordinates we can easily write the differential equations for the trajectories in the familiar form (regard q^i and p_i as functions on M and use $X_H = \{\cdot, H\}$)

$$\dot{q}^i = \{q^i, H\} = \frac{\partial H}{\partial p_i} \tag{C.9a}$$

$$\dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q^i}. \tag{C.9b}$$

We close this appendix with an important remark. In physical applications a symplectic manifold is often given by requiring that the Darboux coordinate functions satisfy

$$\{q^i, q^j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{q^i, p_j\} = \alpha\delta^i_j. \tag{C.10}$$

We immediately see from

$$\{q^i, p_j\} = P^{kl}\partial_l q^i \partial_k p_j = P^{kl}\delta^i_l \delta^{j+n}_k = P^{j+n,i},$$

that the given bracket structure uniquely determines the Poisson tensor (the diagonal blocks have to vanish due to the first two equations in (C.9))

$$P^{mn} = \alpha \begin{pmatrix} \mathbb{0} & -\mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix}, \tag{C.11}$$

and hence also the symplectic form.

APPENDIX D

Mathematical supplements for physicists

We collect in this appendix some basic definitions and theorems that are mandatory for LQG but probably not treated accordingly in most physics curricula. This is not meant to be an appropriate introduction for any of the presented subjects but it should suffice to be able to follow the treatment in the main text. We therefore skip all the proofs (except if they are simple and contribute to the overall understanding) and refer to the exhaustive literature on the subjects. For the representation theory of groups and related subjects [41] is a good source. For functional analytic questions [31] is probably the most common reference.

1. Representation theory

A *representation* of a group G is a homomorphism from G to the group of linear operators on some Hilbert space V. In formulas, $\pi\colon G \to \mathcal{B}(V)$ with $\pi(g_1 g_2) = \pi(g_1)\pi(g_2)$. The dimension of V is called the dimension of the representation. π is called *faithful* if it is injective and *unitary* if every operator $\pi(g)$ is unitary. The trivial representation on a one-dimensional space maps every element to the identity operator. Note that in the case of finite dimensional V the group elements are represented by matrices.

We have the following standard way of constructing new representations out of old ones:

- The dual representation π' on the dual space V' of continuous linear functionals f is defined by
$$(\pi'(g)f)(v) := f(\pi(g^{-1})v).$$
By the Riesz lemma every functional on a Hilbert space can be represented by an inner product with a given element of V, hence $\pi'(g) = (\pi(g^{-1}))^\dagger$ in this identification of V' and V.
- The tensor product of two representations is a representation on the tensor product of V_1 and V_2,
$$(\pi_1 \otimes \pi_2)(g)(v_1 \otimes v_2) := (\pi_1(g)v_1) \otimes (\pi_2(g)v_2).$$
- The direct sum of representations acts on the direct sum of the representation spaces. The obvious definition is
$$\pi(g)(v_1, v_2) := (\pi_1(g)v_1, \pi_2(g)v_2).$$

A subspace $W \subset V$ is called *invariant* if $\pi(g)W \subset W$ for all $g \in G$. Representations which have no invariant subspace (except of course $\{0\}$ and V) are called *irreducible*.

A representation is called completely reducible, if V is an at most countable direct sum of subspaces, $V = \bigoplus_i V_i$, and π is the direct sum of *irreducible* representations on these subspaces.

Two representations are called *equivalent*, if there exists an invertible operator $U\colon V_1 \to V_2$, such that $\pi_2(Ug) = U\pi_1(g)$ for all $g \in G$.

A measure on a locally compact group G is called *left invariant* if for all $g \in G$ and all measurable subsets $S \subset G$ we have $\mu(aS) = \mu(S)$. The definition of a right invariant measure is similar. It can be shown that on a compact finite-dimensional Lie group there exists (up to a multiplicative constant) a unique measure μ_H, called the *Haar measure*, that is left and right invariant. The constant is usually fixed by requiring that μ_H is a probability measure (i.e. $\mu_H(G) = 1$). The Haar measure is also invariant under inversion, $\mu_H(S^{-1}) = \mu_H(S)$ for all $S \subset G$.

Lemma D.1 (Schur). *Suppose we are given two finite-dimensional irreducible representations of the group G. If there exists a linear map $A\colon V_1 \to V_2$ such that $\pi_2(g)A = A\pi_1(g)$ for all $g \in G$, then either $A = 0$ or A is a bijection (hence the representations are then equivalent). In the second case the operator A is also uniquely determined up to a multiplicative constant.*

Proof. For $x \in \ker(A)$ we have $A\pi_1(g)x = \pi_2(g)Ax = 0$, hence $\ker(A)$ is an invariant subspace. Also $\mathrm{Im}(A)$ is invariant, because for $y = Ax \in \mathrm{Im}(A)$ we have $\pi_2(g)y = \pi_2(g)Ax = A\pi_1(g)x$. Since both representations are assumed to be irreducible, kernel and image of A can only be $\{0\}$ or V_1 (resp. V_2). There are therefore only two possible cases. Either $\ker(A) = V_1$ — which means that $A = 0$ — or $\ker(A) = \{0\}$, i.e. A is injective. In the second case $\mathrm{Im}(A) \neq \{0\}$. Therefore $\mathrm{Im}(A) = V_2$ and A is also surjective and therefore a bijection.

Now if we are in the second case, consider any other operator B which also satisfies $\pi_2(g)B = B\pi_1(g)$. Then for every complex λ the operator $A - \lambda B$ also satisfies this equation. In particular we can take λ such that $\det(A - \lambda B) = 0$ (for example take an eigenvalue of $B^{-1}A$). But then $A - \lambda B = 0$ by the part of Schur's lemma that we have already proved. □

Note that in the particular case $\pi_1 = \pi_2 =: \pi$ Schur's lemma strengthens to $A = 0$ or $A = \lambda\,\mathrm{id}_V$, since the operator in the second case is unique up to a constant and id_V is just such an operator.

Schur's lemma allows us to prove our most important theorem for representations of compact Lie groups.

Theorem D.2 (Peter and Weyl). *Let $(\pi_j)_{j\in J}$ be a family of finite-dimensional, unitary, irreducible representations of the compact Lie group G, where from* every *equivalence class of such representations precisely one representative has been chosen. Then the set*

$(b_{jmn})_{j\in J, m,n\in\{1,\dots,d_j\}}$ *of matrix element functions, multiplied by the square root of the dimension of the representation,*

$$b_{jmn}\colon g \mapsto \sqrt{d_j}(\pi_j(g))_{mn},$$

is a complete orthonormal basis of the space $L_2(G, d\mu_H)$.

Proof.

(1) Orthonormality

For an arbitrary linear map $M\colon V_j \to V_{j'}$ (i.e. a matrix), consider

$$A_M := \int_G d\mu_H(g)(\pi_{j'}(g))^\dagger M\pi_j(g). \tag{D.1}$$

A_M is obviously also a linear map from V_j to $V_{j'}$ and satisfies

$$\begin{aligned} A_M\pi_j(g) &= \int_G d\mu_H(g')(\pi_{j'}(g'))^\dagger M\pi_j(g')\pi_j(g) \\ &= \int_G d\mu_H(g')(\pi_{j'}(g'))^\dagger M\pi_j(g'g) \\ &= \int_G d\mu_H(g')(\pi_{j'}(g'g^{-1}))^\dagger M\pi_j(g') \\ &= \pi_{j'}(g)\int_G d\mu_H(g')(\pi_{j'}(g'))^\dagger M\pi_j(g') = \pi_{j'}(g)A_M, \end{aligned}$$

since by right invariance of the Haar measure we can act with $R_{g^{-1}}$ on the argument of the integrand and where we used unitarity of the representation matrices. We see that A_M satisfies the assumptions of Schur's lemma and conclude that $A_M = 0$ for $j \neq j'$, because then the representations are inequivalent. For $j = j'$ Schur's lemma states that $A_M = \lambda_M \,\mathrm{id}_V$. The value of λ_M is found by taking the trace of (D.1), giving $\lambda_M d_j = \mathrm{tr}(M)$. Taking $(M)_{kl} = \delta_{km'}\delta_{lm}$ then gives orthonormality,

$$\begin{aligned} (b_{j'm'n'}, b_{jmn}) &= \sqrt{d_j d_{j'}}\int_G d\mu_H \overline{(\pi_{j'})}_{m'n'}(\pi_j)_{mn} \\ &= \sqrt{d_j d_{j'}}\int_G d\mu_H \overline{(\pi_{j'})}_{kn'}(M)_{kl}(\pi_j)_{ln} \\ &= \sqrt{d_j d_{j'}}(A_M)_{n'n} = \sqrt{d_j d_{j'}}\delta_{jj'}\lambda_M(\mathrm{id}_V)_{n'n} \\ &= \delta_{jj'}\delta_{mm'}\delta_{nn'}. \end{aligned}$$

(2) Completeness

To show completeness one verifies that $\mathrm{span}\,\{b_{jmn}\}$ satisfies all the assumptions of the Stone-Weierstrass theorem (just like in the proof of lemma 3.14) to see that $\mathrm{span}\,\{b_{jmn}\}$ is dense in $C(G)$. Completeness is then immediate because $C(G)$ is in turn dense in $L_2(G, d\mu_H)$ (see e.g. theorem 3.14 in [34]).

□

1.1. Representations of $SU(2)$. The representation theory of $SU(2)$ should be familiar from the treatment of angular momentum in quantum mechanics. But since the connection to group representations is often not made, we recall the basics here.

The irreducible, unitary representations of $SU(2)$ are labeled by a half-integral number $j \in \{0, \frac{1}{2}, 1, \dots\}$. The most common representations are the trivial one for $j = 0$, and the fundamental representation $j = \frac{1}{2}$, which is defined by the natural action of complex, unitary 2×2-matrices with unit determinant on vectors in $\mathbb{C}^2$ (these vectors are called spinors). The representation $j = 1$ is called the adjoint representation.

All other irreducible representations can be constructed from the fundamental one. The representation space for the spin-j representation is the space of completely symmetric spinors with $n = 2j$ indices. The action of the $SU(2)$-matrix $U^A{}_B$ is defined by letting this matrix act on all the indices,

$$\psi^{A_1 \dots A_n} \mapsto U^{A_1}{}_{A'_1} U^{A_n}{}_{A'_n} \psi^{A'_1 \dots A'_n}.$$

The dimension of the spin-j representation is therefore $2j + 1$. For simplicity we will often collect the indices $A_1 \dots A_{2j}$ of completely symmetric spinors in a single index $m = A_1 + \dots + A_{2j}$ and write $\psi^m = \psi^{(A_1 \dots A_{2j})}$. The representation matrices are then denoted by $(\rho_j(g))^m{}_n := U^{(A_1}{}_{(B_1} \dots U^{A_{2j})}{}_{B_{2j})}$.[1]

The Clebsh-Gordan theorem states that the tensor product of two irreducible representations is completely reducible and decomposes into the direct sum of irreducible representations, $\pi_{j_1} \otimes \pi_{j_2} = \pi_{|j_1 - j_2|} \oplus \pi_{|j_1 - j_2|+1} \oplus \dots \oplus \pi_{j_1 + j_2}$. For the elements of the representation matrices this means that $(\rho_{j_1})_{m_1 n_1} (\rho_{j_2})_{m_2 n_2} = \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m,n} c_{jmn} (\rho_j)_{mn}$. To see this, notice that the LHS can be interpreted as a matrix element in the tensor product. For the RHS, remember that the representation matrix of a direct sum takes block-diagonal form in a suitable basis, the blocks being the representation matrices of the summands (the coefficients c_{jmn} take into account that in general we have to perform a change of basis to achieve block-diagonal form — their explicit values are not of importance for our purposes).

Finally we introduce the notion of an *intertwiner* between representations $j_1, j_2, \dots, j_N$. Their tensor product is completely reducible and can therefore be decomposed into a direct sum of irreducible components. An element of the subspace which belongs to the trivial representation is called an intertwiner. In other words, elements of the tensor product of the representation spaces are tensors $v^{m_1 \dots m_N}$, and intertwiners are tensors invariant under the action of $SU(2)$,

$$\rho_j(U)^{m_1}{}_{n_1} \dots \rho_j(U)^{m_N}{}_{n_N} i^{n_1 \dots n_N} = i^{m_1 \dots m_N}. \tag{D.2}$$

[1]Another possibility is to choose a basis $e_i^{A_1 \dots A_{2j}}$ in space of completely symmetric spinors and let the index m denote the components in this basis. For example for the adjoint representation $j = 1$, a basis is provided by $e_i^{AB} := \frac{1}{\sqrt{2}} \sigma_i^{AB} := \frac{1}{\sqrt{2}} \sigma_i{}^A{}_C \varepsilon^{CB}$ with the Pauli matrices $\sigma_i{}^A{}_C$.

In LQG we actually need intertwiners between representations $j_1, j_2, \ldots, j_k$ and the *dual* representations of $j_{k+1}, \ldots, j_N$. We can however construct these out of the intertwiners $i^{m_1 \ldots m_N}$ by lowering indices. To lower the index m_1 for example, expand it into indices of a completely symmetric tensor, $A_1 \ldots A_{2j_1}$ and then lower every index with the completely antisymmetric symbol ε_{AB}, defined with $\varepsilon_{01} = \varepsilon^{01} = 1$. Here is the explicit invariance property of these intertwiners,

$$\rho_j(U)^{m_1}{}_{m_1'} \ldots \rho_j(U)^{m_k}{}_{m_k'} \rho_j(U^{-1})^{n_{k+1}'}{}_{n_{k+1}} \ldots \rho_j(U^{-1})^{n_N'}{}_{n_N} i^{m_1' \ldots m_k'}{}_{n_{k+1}' n_N'} = i^{m_1 \ldots m_k}{}_{n_{k+1} n_N}. \tag{D.3}$$

We will prove this assertion only for an intertwiner between two fundamental representations ($j = \frac{1}{2}$) — the general case follows easily since all other representations are constructed from the fundamental one (see above). First observe that ε_{AB} and ε^{AB} are invariant under the action of $SU(2)$,

$$U^A{}_C U^B{}_D \varepsilon^{CD} = \varepsilon^{AB}. \tag{D.4}$$

Since both sides are antisymmetric 2×2-tensors, equality can be checked by contraction with $\frac{1}{2}\varepsilon_{AB}$, which gives the only independent component (remember expression (N.20) for the determinant),

$$1 = \det U = \frac{1}{2}\varepsilon_{AB} U^A{}_C U^B{}_D \varepsilon^{CD} = \frac{1}{2}\varepsilon^{AB}\varepsilon_{AB} = 1.$$

Equation (D.4) tells us that the fundamental representation is equivalent to its dual representation $\psi_A \mapsto (U^\dagger)^B{}_A \psi_B$ (this is the reason why we do not pull indices with δ_{AB}). Namely, application of $(U^{-1})^E{}_A$ gives

$$U^B{}_D \varepsilon^{ED} \psi_E = \varepsilon^{AB} (U^\dagger)^E{}_A \psi_E, \tag{D.5}$$

or written for ε_{AB},

$$\varepsilon_{BC} U^B{}_F = (U^\dagger)^E{}_C \varepsilon_{EF}. \tag{D.6}$$

With these preparations it is easy to see that a given intertwiner i^{AB} leads to an intertwiner $i_A{}^B := \varepsilon_{AA'} i^{A'B}$,

$$(U^\dagger)^{A'}{}_A U^B{}_{B'} i_{A'}{}^{B'} = (U^\dagger)^{A'}{}_A U^B{}_{B'} \varepsilon_{A'C} i^{CB'} = \varepsilon_{A'A} U^{A'}{}_C U^B{}_{B'} i^{CB'} = \varepsilon_{A'A} i^{A'B} = i_A{}^B.$$

The trivial representation can appear multiple times in the decomposition into irreducibles. In this case we choose a basis of intertwiners satisfying the orthonormality condition

$$\overline{i_i^{m_1 \ldots m_N}} i_{j, m_1 \ldots m_N} = \delta_{ij}. \tag{D.7}$$

2. Miscellaneous

The standard reference for physicists concerning functional analytic questions is [31]. This book also contains the proofs of the following theorems.

For the Stone-Weierstrass theorem we introduce the following terminology. Let $\mathcal{A}$ be an algebra of functions on the set K. $\mathcal{A}$ is said to *separate the points* of K, iff for every

$x_1, x_2 \in K$ with $x_1 \neq x_2$ there is a function $f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$. $\mathcal{A}$ is said to be *nowhere vanishing* iff for every $x \in K$ there is a function $f \in \mathcal{A}$ with $f(x) \neq 0$.

THEOREM D.3 (Stone-Weierstrass). *Let $\mathcal{A} \subset C(K, \mathbb{R})$ be an algebra of continuous functions on a compact space K. If $\mathcal{A}$ is nowhere vanishing and separates the points of K, then $\mathcal{A}$ is dense in the algebra $C(K, \mathbb{R})$ of* all *continuous functions on K.*

The theorem holds also for an algebra of complex-valued functions, as long as this algebra is also closed under complex conjugation (for every $f \in \mathcal{A}$, $\overline{f}$ is also an element of $\mathcal{A}$).

A linear functional Λ on the algebra $C(K)$ is called positive, iff $\Lambda(f) \geq 0$ for all positive functions $f \in C(K)$. We will be working with the following special case of the Riesz representation theorem.

THEOREM D.4 (Riesz). *Let K be a compact Hausdorff space and $\Lambda\colon C(K) \to \mathbb{C}$ a positive linear functional on the space of continuous functions on K with $\Lambda(1) = 1$. Then there is a unique, regular Borel probability measure μ that represents Λ in the sense that*

$$\Lambda(f) = \int_K f d\mu \quad \forall f \in C(K).$$

Bibliography

[1] Vladimir I. Arnold. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate texts in mathematics*. Springer, 2nd edition, 1989.

[2] Richard L. Arnowitt, Stanley Deser, and Charles W. Misner. The dynamics of general relativity. In Louis Witten, editor, *Gravitation: an introduction to current research*, chapter 7, pages 227–265. Wiley, 1962, gr-qc/0405109.

[3] Richard L. Arnowitt, Stanley Deser, and Charles W. Misner. Republication of: The dynamics of general relativity. *General Relativity and Gravitation*, 40:1997–2027, September 2008, gr-qc/0405109.

[4] Abhay Ashtekar. *Lectures on nonperturbative canonical gravity*, volume 6 of *Advanced series in astrophysics and cosmology*. World Scientific, Singapore, Singapore, 1991.

[5] Abhay Ashtekar, Martin Bojowald, and Jerzy Lewandowski. Mathematical structure of loop quantum cosmology. *Adv. Theor. Math. Phys.*, 7:233–268, 2003, gr-qc/0304074.

[6] Abhay Ashtekar and Jerzy Lewandowski. Quantum theory of geometry. II: Volume operators. *Adv. Theor. Math. Phys.*, 1:388–429, 1998, gr-qc/9711031.

[7] Abhay Ashtekar, Tomasz Pawlowski, and Parampreet Singh. Quantum Nature of the Big Bang: An Analytical and Numerical Investigation I, April 2006, gr-qc/0604013.

[8] Abhay Ashtekar, Tomasz Pawlowski, and Parampreet Singh. Quantum nature of the big bang: Improved dynamics. August 2006, gr-qc/0607039.

[9] Herbert Balasin. Geometrie und Gravitation II. Lecture at Vienna University of Technology, 2009-2010.

[10] Martin Bojowald. Loop quantum cosmology. I: Kinematics. *Class. Quant. Grav.*, 17:1489–1508, 2000, gr-qc/9910103.

[11] Martin Bojowald. Loop quantum cosmology. II: Volume operators. *Class. Quant. Grav.*, 17:1509–1526, 2000, gr-qc/9910104.

[12] Martin Bojowald. Absence of singularity in loop quantum cosmology. *Phys. Rev. Lett.*, 86:5227–5230, 2001, gr-qc/0102069.

[13] Martin Bojowald. Loop quantum cosmology. III: Wheeler-DeWitt operators. *Class. Quant. Grav.*, 18:1055–1070, 2001, gr-qc/0008052.

[14] Martin Bojowald. Loop quantum cosmology. IV: Discrete time evolution. *Class. Quant. Grav.*, 18:1071–1088, 2001, gr-qc/0008053.

[15] Martin Bojowald. The inverse scale factor in isotropic quantum geometry. *Phys. Rev.*, D64:084018, 2001, gr-qc/0105067.

[16] Martin Bojowald. Isotropic loop quantum cosmology. *Class. Quant. Grav.*, 19:2717–2742, 2002, gr-qc/0202077.

[17] Martin Bojowald and Hans A. Kastrup. Quantum symmetry reduction for diffeomorphism invariant theories of connections. *Class. Quant. Grav.*, 17:3009–3043, 2000, hep-th/9907042.

[18] Othmar Brodbeck. On symmetric gauge fields for arbitrary gauge and symmetry groups. *Helv. Phys. Acta*, 69:321–324, 1996, gr-qc/9610024.

[19] Sean M. Carroll. *Spacetime and geometry: An introduction to general relativity.* Addison-Wesley, San Francisco, USA, 2004.

[20] Yvonne Choquet-Bruhat and Cécile DeWitt-Morette. *Analysis, Manifolds and Physics, Part I.* North Holland, Amsterdam, 1996.

[21] Bryce S. DeWitt. Quantum theory of gravity. I. The canonical theory. *Phys. Rev.*, 160:1113–1148, 1967.

[22] Bryce S. DeWitt. Quantum theory of gravity. II. The manifestly covariant theory. *Phys. Rev.*, 162:1195–1239, 1967.

[23] Bryce S. DeWitt. Quantum theory of gravity. III. Applications of the covariant theory. *Phys. Rev.*, 162:1239–1256, 1967.

[24] Paul A. M. Dirac. *Lectures on Quantum Mechanics.* Dover Publications, March 2001.

[25] Stephen W. Hawking and George F. R. Ellis. *The Large scale structure of space-time.* Cambridge University Press, Cambridge, UK, 1973.

[26] Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems.* Princeton University Press, August 1994.

[27] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry, Volume I.* Wiley-Interscience, February 1996.

[28] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation.* W. H. Freeman and Company, San Francisco, USA, 1973.

[29] Roger Penrose. Theory of quantized directions, unpublished manuscript; Angular momentum: an approach to combinatorial spacetime. In T. Bastin, editor, *"Quantum Theory and Beyond"*, pages 151–180. Cambridge University Press, Cambridge, 1971.

[30] Michael E. Peskin and Dan V. Schroeder. *An Introduction To Quantum Field Theory (Frontiers in Physics).* Westview Press, October 1995.

[31] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics*, volume 1-4. Academic Press, Boston, 1980.

[32] Carlo Rovelli. *Quantum gravity.* Cambridge University Press, Cambridge, UK, 2004.

[33] Carlo Rovelli and Lee Smolin. Discreteness of area and volume in quantum gravity. *Nucl.Phys.*, B442:593–622, 1995, gr-qc/9411005.

[34] Walter Rudin. *Real and Complex Analysis.* McGraw-Hill Science/Engineering/Math, 3rd edition, 1986.

[35] Walter Rudin. *Functional Analysis.* McGraw-Hill Science/Engineering/Math, 2 edition, January 1991.

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

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

[38] Thomas Thiemann. Lectures on loop quantum gravity. *Lect. Notes Phys.*, 631:41–135, 2003, gr-qc/0210094.

[39] Thomas Thiemann. The Phoenix project: Master constraint programme for loop quantum gravity. *Class. Quant. Grav.*, 23:2211–2248, 2006, gr-qc/0305080.

[40] Thomas Thiemann. *Modern canonical quantum general relativity.* Cambridge University Press, Cambridge, UK, 2007.

[41] Naum Y. Vilenkin. *Special Functions and the Theory of Group Representations.* American Mathematical Society, Providence, Rhode Island, 1968.

[42] Robert M. Wald. *General Relativity.* Chicago University Press, Chicago, 1984.

[43] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*, volume 94 of *Graduate texts in mathematics.* Springer, 1983.

[44] David L. Wiltshire. Cosmic clocks, cosmic variance and cosmic averages. *New Journal of Physics*, 9(10):377, January 2007.