Introduction to Modern Canonical Quantum General Relativity

T. Thiemann[1]
MPI f. Gravitationsphysik, Albert-Einstein-Institut,
Am Mühlenberg 1, 14476 Golm near Potsdam, Germany

Preprint AEI-2001-119

[1]thiemann@aei.mpg.de

Abstract

This is an introduction to the by now fifteen years old research field of canonical quantum general relativity, sometimes called “loop quantum gravity”. The term “modern” in the title refers to the fact that the quantum theory is based on formulating classical general relativity as a theory of connections rather than metrics as compared to in original version due to Arnowitt, Deser and Misner.

Canonical quantum general relativity is an attempt to define a mathematically rigorous, non-perturbative, background independent theory of Lorentzian quantum gravity in four spacetime dimensions in the continuum. As such it differs considerably from perturbative ansätze. It provides a *unified* theory of all interactions in the sense that all interactions transform under a common gauge group: The four-dimensional diffeomorphism group of the underlying differental manifold which is maximally broken in perturbative approaches. The approach is minimal in that one simply analyzes the logical consequences of combining the principles of general relativity with the principles of quantum mechanics. As a consequence, no extra dimensions, no corresponding Kaluza-Klein compactifications, no supersymmetry and its associated phenomenology – compatible spontaneous breaking at low energies, seem to be necessary. On the other hand, no explanation for the particle content and the dimension of the universe is provided by the theory.

The requirement to preserve background independence has lead to new, fascinating mathematical structures which one does not see in perturbative approaches, e.g. a fundamental discreteness of spacetime seems to be a prediction of the theory which is a first substantial evidence for a theory in which the gravitational field acts as a natural UV cut-off.

An effort has been made to provide a self-contained exposition at the appropriate level of rigour which at the same time is accessible to graduate students with only basic knowledge of general relativity and quantum field theory on Minkowski space. To be useful, not all facets of the field have been covered, however, guides to further reading and a detailed bibliography is included. This report is submitted to the on-line journal “Living Reviews” and is thus subject to being updated on at least a bi-annual basis.

Contents

Introduction

This report tries to give a status overview of the field of modern canonical quantum general relativity, sometimes called "loop quantum gravity". The term "modern" accounts for the fact that this is a "connection dynamics" formulation of Einstein's theory, rather than the original "geometrodynamics" formulation due to Arnowitt, Deser and Misner. As this is an article submitted to the on-line journal "Living Reviews, the report will be updated on an at least bi-annual basis.

The field of modern canonical quantum general relativity was born in 1986 and since then an order of 500 research papers closely related to the subject have been published. Pivotal structures of the theory are scattered over an order of 100 research papers, reports, proceedings and books, issues which were believed to be essential initially turned out to be negligible later on and vice versa. These facts make it hard for the beginner to go quickly to the frontier of original research. The present report aims at giving beginners a guideline, a kind of *geodesic* through the literature suggesting in which order to read a minimal number of papers in order to understand the foundations. The report should be accessible to students at the graduate level (say European students in their sixth or seventh semester) with only basic prior knowledge of general relativity and quantum field theory on Minkowski space. However, the target audience does not consist of the light hearted readers who wish to get just some rough idea of what the field is all about, those readers are advised to study the excellent review articles [1, 2, 3]. Rather, we typically have in mind the serious student who wants to gain thorough understanding of the foundations and to pass quickly to the frontier of active research. As financial support over an extended period of time is a serious issue for almost all graduate students in the world, time is pressing and it is important that one does not waste too much time in learning the basics. We have therefore decided to select only a minimal amount of material, just enough in order to reach a firm understanding of the foundations, which on the other hand is presented in great detail so that one does not have to read much besides this report for this purpose. Since some of the mathematics that is needed maybe unfamiliar to the younger students we have also included a large mathematical "appendix" that serves to fill in some of the necessary background. Remembering too freshly still our own experience how annoying and time consuming it can be to collect papers, to compare and streamline notations, to adapt numerical coefficient conventions etc. this should also help to avoid unnecessary confusions and time delays.

The number of people working in the field of canonical quantum general relativity is of the order of 10^2 (including students and post-docs) which is unfortunately quite small, when compared, for instance, with the size of the string theory community (of the order 10^3) and it is one of the aims of this report to attract more young researchers to dive into this alternative, fascinating research subject. Here is a, to the best knowledge of the author, complete list of locations where research in canonical quantum general relativity (and related) is, at least partly, currently performed *and funded* together with the names of main contact persons in alphabetical order and their main current research directions. The locations are listed alphabetically by continent and within one continent geographically by country from north to south (p=permanent position, n=non-permanent position, we only list post-docs among the non-permanent staff):

A) **Central America**

1. *Universidad Autonoma Metropolitana Itztapalapa, Mexico City, Mexico*
   Hugo Morales-Tecotl (p): semiclassical quantum gravity
2. *Universidad Nacionale Autonoma de Mexico, Mexico City, Mexico*

Alexandro Corichi (p): semiclassical quantum gravity, isolated horizon quantum black holes;
Michael Ryan (p) : mini(midi)superspace models;
Jose-Antonio Zapata (p) : semiclassical quantum gravity, spin foams

B) **North America**

3. *University of Alberta, Edmonton, Canada*

   Stephen Fairhurst (n): isolated horizon quantum black holes;
   Viquar Husain (p): mini(midi)superspace and integrable models;
   Don Page (p): mini(midi)superspace models

4. *Perimeter Institute for Theoretical Physics and University of Waterloo, Waterloo, Ontario, Canada*

   Olaf Dreyer (n): isolated horizon quantum black holes;
   Fotini Markopoulou (p): causal spin foams, renormalization;
   Lee Smolin (p): connections between loop quantum gravity and string theory, causal quantum dynamics, spin foam models;
   O. Winkler (n): semi-classical quantum gravity

5. *Syracuse University, Syracuse, NY, USA*

   Donald Marolf (p): refined algebraic quantization, general relativistic aspects of string theory

6. *Center for Gravitational Physics and Geometry, The Pennsylvania State University at University Park, PA, USA*

   Abhay Ashtekar (p): isolated horizon quantum black holes, semi-classical quantum gravity;
   Martin Bojowald (n): mini(midi)superspace models, quantum dynamics;
   Amit Ghosh (n): isolated horizon black holes;
   Roger Penrose[1] (p): twistor theory, fundamental issues;
   Alexandro Perez (n): spin foam models

7. *University of Utah, Salt Lake City, UT, USA*

   Christopher Beetle (n): covariant formulation;
   Karel Kuchař (p): mini – and midisuperspace models, covariant formulation

8. *University of Maryland, College Park, MD, USA*

   Ted Jacobson (p): classical actions, (quantum) black hole physics

9. *Kansas State Unviversity, Kansas City, KS, USA*

   Louis Crane (p): state sum models, spin foam models;
   David Yetter(p): state sum models, spin foam models

10. *Utah State University, Logan, UT, USA*

    Charles Torre (p): mini(midi)superspace models, fundamental issues

11. *University of California, Riverside, CA, USA*

    John Baez (p): isolated horizon quantum black holes, spin foam models

[1] Distinguished Visiting Professor from the University of Oxford, Oxford, UK

12. *University of California and Institute of Theoretical Physics, Santa Barbara, CA, USA*
    Jim Hartle (p): connection between consistent histories – and canonical approach;
    Kirill Krasnov (n) : isolated horizon quantum black holes, spin foam models, aspects of string theory
13. *University of Mississippi, Oxford, MS, USA*
    Luca Bombelli (p): semiclassical quantum gravity
14. *Lousiana State University, Baton Rouge, LA, USA*
    Jorge Pullin (p): quantum dynamics, Dirac observables, semiclassical quantum gravity

C) **South America**

15. *Universidad de la Republica, Montevideo, Uruguay*
    Rodolfo Gambini (p): quantum dynamics, Dirac observables;
    Jorge Griego (p): quantum dynamics;
    Michael Reisenberger (p): dynamical lattice formulations, spin foam models
16. *Centro de Estudios Cientificos, Valdivia, Chile*
    Claudio Teitelboim: mini(midi)superspace models, quantization of gauge systems;
    Andres Gomberoff (p): refined algebraic quantization, general relativistic aspects of string theory

D) **Asia**

17. *Raman Research Institute, Bangalore, India*
    Joseph Samuel (p): classical formulation;
    Madhavan Varadarajan (p): semiclassical quantum gravity

E) **Europe**

18. *The Niels Bohr Institute, Copenhagen, Denmark*
    Jan Ambjorn (p): path integral formulation (dynamical triangulations);
    Matthias Arnsdorf (n): semiclassical quantum gravity, connections with string theory
19. *University of Nottingham, Nottingham, UK*
    John Barrett (p): spin foam models, state sum models;
    Jorma Louko (p): mini(midi)superspace models, general relativistic aspects of string theory
20. *Albert – Einstein – Institut, Golm near Potsdam, Germany*
    Hermann Nicolai (p): supergravity, superstring theory, connections between canonical quantum gravity and M – Theory;
    Thomas Thiemann (p): quantum dynamics, semi-classical analysis
21. *University of Warsaw, Warsaw, Poland*
    Jurek Lewandowski (p): isolated horizon quantum black holes, semi-classical quantum gravity
22. *Utrecht University and Spinoza Institute, Utrecht, The Netherlands*
    Gerhard 't Hooft (p): quantum black hole physics, fundamental issues;
    Renate Loll (p): quantum geometry, path integral formulation (dynamical triangulations)

23. *Cambridge University, Cambridge, UK*
    Ruth Williams (p): spin foam models, state sum models
24. *Imperial College, London, UK*
    Chris Isham (p): fundamental issues (topos theory)
25. *Université Libre de Bruxelles, Bruxelles, Belgium*
    Marc Henneaux (p): BRST analysis, quantization of gauge systems, general relativistic aspects of supergravity and string theory
26. *Ecole Normale Supérieure, Paris, France*
    Bernard Julia (p): canonical quantization of supergravity theories, Noether charges
27. *Universität Wien, Wien, Austria*
    Peter Aichelburg (p): general relativistic aspects of supergravity and string theory
28. *Universität Bern, Bern, Switzerland*
    Peter Hajíček (p): midi(mini)superspace models
29. *Ecole Normale Supérieur, Lyon, France*
    Laurent Freidel (p) : spin foam models
30. *Universita di Torino, Torino, Italy*
    Jeanette Nelson (p): mini(midi)superspace models, Regge calculus
31. *Universita di Parma, Parma, Italy*
    Roberto de Pietri (n) : quantum dynamics, spin foam models
32. *Université de Marseille, Luminy, France*
    Carlo Rovelli (p): quantum dynamics, Dirac observables, spin foam models
33. *Instituto de Matematicas y Fisica Fundamental, Madrid, Spain*
    Guillermo Mena-Maguán (p): mini(midi)superspace models, Euclidean versus Lorentzian formulation
34. *Universidad Europea, Madrid, Spain*
    Fernando Barbero (p): classical actions
35. *Instituto Superior Tecnico, Lisboa, Portugal*
    José Mourão (p): mathematical framework
36. *Universidad do Algarve, Faro, Portugal*
    Nenad Manojlovic (p): mini(midi)superspaces, integrable models

All these places are definitely worthwhile applying to for graduate studies or post-doc positions. Notice that we have listed only those researchers that are at least partly involved or interested in quantum general relativity research and only those aspects of their work that touch on quantum gravity. For instance, the Albert – Einstein – Institute is currently the largest (by number of members and budget) institute in the world that focusses on gravitational physics, consisting of altogether four divisions: Quantum Gravity and Unified Theories with focus on M – Theory (director: Hermann Nicolai), Astrophysics (director: Bernard Schutz), Mathematical General Relativity (director: Gerhard Huisken), Gravitational Wave Detection (director: Karsten Danzmann). Institutes of similar sizes are the Center for Gravitational Physics and Geometry, the Perimeter Institute for Theoretical Physics and the Institute for Theoretical Physics at Santa Barbara. Unfortunately the University of

Pittsburgh, Pittsburgh, PA, USA (Ted Newman (p): null surface formulation, fundamental issues) no longer appears in the above list since quantum gravitational research is no longer funded there.

This report is organized as follows :

Before diving into the subject, the next subsection motivates the search for a quantum theory of gravity, points out what the essential problems are that one has to deal with, lists the possible approaches and their respective strengths and weaknesses and finally motivates our choice to study canonical quantum general relativity. We also list our notation and conventions.

We then approach the main text of the report which is subdivided into three parts:

The first part contains the foundations of the theory, that is, results which are physically and mathematically robust. Thus we will study in detail a) the classical canonical formulation of general relativity in terms of connection variables, b) the general programme of canonical quantization, c) the application of this programme to general relativity in terms of connections and the resulting Hilbert space structures, d) the proof that the Hilbert space found implements the correct quantum kinematics (that is, it represents the correct commutation relations and supports the kinematical constraints of the theory) and finally e) the kinematical geometrical operators which measure for instance areas of (coordinate) surfaces. We also sketch how spectral properties of these kinematical operators extend to their physical (dynamical) counterparts in the presence of matter.

The second part discusses the main, current research directions within canonical quantum general relativity and describes their respective status. These results are less robust and still, at least partially, in flow. Thus we outline in detail a) the implementation of the quantum dynamics or **Quantum Einstein Equantions** (also known by "Wheeler – DeWitt Equation"), b) the coupling of standard quantum matter, c) the semiclassical analysis necessary in order to verify whether the theory constructed is indeed a quantization of general relativity, d) the path integral formulation of the theory (also called spin foam formulation), e) quantum black hole physics and finally f) the possible links between string theory and canonical quantum general relativity. This part closes with a section in which we list a selected number of open and fascinating research problems.

Finally, in the third part we provide some mainly mathematical background material. Thus we give elementary but fairly extensive introductions to elements of a) the Dirac algorithm for dealing with theories with constraints, b) the theory of fibre bundles, c) general topology, d) Gel'fand theory for Abelean C^*-algebras, e) measure theory, f) the GNS construction and g) refined algebraic quantization (RAQ).

Acknowledgements

My thanks go to Theresa Velden, for a long time managing director of the on-line journal "Living Reviews", for continuously encouraging me to finally finish this review.

Defining Quantum Gravity

In the first subsection of this section we explain why the problem of quantum gravity cannot be ignored in nowadays physics, even though the available accelerator energies lie way beyond the Planck scale. Then we define what a quantum theory of gravity and all interactions is widely expected to achieve and point out the two main directions of research divided into the perturbative and non-perturbative approaches. In the third subsection we describe these approaches in more detail and finally in the fourth motivate our choice to do canonical quantum general relativity as opposed to other approaches.

Why Quantum Gravity in the 21st Century ?

It is often argued that quantum gravity is not relevant for the physics of this century because in our most powerful accelerator, the LHC to be working in 2005, we obtain energies of the order of a few 10^3 GeV while the energy scale at which quantum gravity is believed to become important is the Planck energy of 10^{19} GeV. While that is true, it is false that nature does not equip us with particles of energies much beyond the TeV scale, there are astrophysical particles with energy of a fist stroke and the next generation of particle microscopes is therefore not going to be built on the surface of Earth any more but in its orbit. Moreover, as we will describe in this report in more detail, even with TeV energy scales it might be possible to see quantum gravity effects in the close future.

But even apart from these purely experimental considerations, there are good theoretical reasons for studying quantum gravity. To see why, let us summarize our current understanding of the fundamental interactions :
Ashamingly, the only quantum fields that we fully understand to date in four dimensions are *free quantum fields on four-dimensional Minkowski space.* Formulated more provocatively:

In four dimensions we only understand an (infinite) collection of uncoupled harmonic oscillators on Minkowski space !

In order to leave the domain of these rather trivial and unphysical quantum field theories, physicists have developed two techniques : perturbation theory and quantum field theory on curved backgrounds. This means the following :
With respect to accelerator experiments, the most important processes are scattering amplitudes between particles. One can *formally* write down a unitary operator that accounts for the scattering interaction between particles and which maps between the well-understood free quantum field Hilbert spaces in the far past and future. Famously, by Haags theorem [4], whenver that operator is really unitary, there is no interaction and if it is not unitary, then it is ill-defined. In fact, one can only define the operator perturbatively by writing down the formal power expansion in terms of the generator of the would-be unitary transformation between the free quantum field theory Hilbert spaces. The resulting series is divergent order by order but if the theory is “renormalizable” then one can make these orders artificially finite by a regularization and renormalization procedure with, however, no control on convergence of the resulting series. Despite these drawbacks, this recipe has worked very well so far, at least for the electroweak interaction.

Until now, all we have said applies only to free (or perturbatively interacting) quantum fields on Minkowski spacetime for which the so-called Wightman axioms [4] can be verified. Let us summarize them for the case of a scalar field in $(D+1)-$dimensional Minkowski space:

W1 *Representation*
There exists a unitary and continuous representation $U: \boldsymbol{\mathcal{P}} \to \mathcal{B}(\mathcal{H})$ of the **Poincaré group** $\boldsymbol{\mathcal{P}}$ on a Hilbert space $\mathcal{H}$.

W2 *Spectral Condition*
The momentum operators $\boldsymbol{P^\mu}$ have spectrum in the forward lightcone:
$\boldsymbol{\eta_{\mu\nu} P^\mu P^\nu} \leq \mathbf{0};\ \boldsymbol{P^0} \geq \mathbf{0}$.

W3 *Vacuum*
There is a unique **Poincaré** invariant vacuum state $U(\boldsymbol{p})\boldsymbol{\Omega} = \boldsymbol{\Omega}$ for all $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}$.

W4 *Covariance*
Consider the smeared field operator valued tempered distributions $\phi(f) = \int_{\mathbb{R}^{\boldsymbol{D+1}}} d^{D+1}x \phi(x) f(x)$ where $f \in \mathcal{S}(\mathbb{R}^{\boldsymbol{D+1}})$ is a test function of rapid decrease. Then finite linear combinations of the form $\phi(f_1)..\phi(f_N)\boldsymbol{\Omega}$ lie dense in $\mathcal{H}$ (that is, $\boldsymbol{\Omega}$ is a cyclic vector) and $U(p)\phi(f)U(p)^{-1} = \phi(f \circ p)$ for any $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}$.

W5 *Locality (Causality)*
Suppose that the supports (the set of points where a function is different from zero) of f, f' are *spacelike separated* (that is, the points of their supports cannot be connected by a non-spacelike curve) then $\phi(f), \phi(f') = 0$.

The most important objects in this list are those that are highlighted in bold face letters: The fixed, non-dynamical Minkowski background metric $\boldsymbol{\eta}$ with its well-defined causal structure, its Poincaré symmetry group $\boldsymbol{\mathcal{P}}$, the associated representation $U(\boldsymbol{p})$ of its elements, the invariant vacuum state $\boldsymbol{\Omega}$ and finally the fixed, non-dynamical topological, differentiable manifold $\mathbb{R}^{\boldsymbol{D+1}}$. Thus the Wightman axioms assume the existence of a non-dynamical, Minkowski background metric which implies that we have a preferred notion of causality (or locality) and its symmetry group, the Poincaré group from which one builds the usual Fock Hilbert spaces of the free fields. We see that the whole structure of the theory is heavily based on the existence of these objects which come with a fixed, non-dynamical background metric on a fixed, non-dynamical topological and differentiable manifold.

For a general background spacetime, things are already under much less control: We still have a notion of causality (locality) but generically no symmetry group any longer and thus there is no obvious generalization of the Wightman axioms and no natural perturbative Fock Hilbert space any longer. These obstacles can partly be overcome by the methods of algebraic quantum field theory [5] and the so-called microlocal analysis [6] (in which the locality axiom is taken care of pointwise rather than globally) which recently have also been employed to develop perturbation theory on arbitrary background spacetimes [7] by invoking the mathematically more rigorous implementation of the renormalization programme developed by Epstein and Glaser in which no divergent expressions ever appear at least order by order (see, e.g., [8]).

However, the whole framework of ordinary quantum field theory breaks down once we make the gravitational field (and the differentiable manifold) dynamical, once there is no background metric any longer !

Combining these issues, one can say that we have a working understanding of scattering processes between elementary particles in arbitrary spacetimes as long as the backreaction of matter on geometry can be neglected and that the coupling constant between non-gravitational interactions is small enough (with QCD being an important exception) since then the classical Einstein equation, which says that curvature of geometry is proportional to the stress energy of matter, can be approximately solved by neglecting matter altogether. Thus, for this set-up, it seems fully sufficient to have only a classical theory of general relativity and perturbative quantum field theory on curved spacetimes.

From a fundamental point of view, however, this state of affairs is unsatisfactory for many reasons among which we have the following:

i) *Classical Geometry – Quantum Matter Inconsistency*
At a fundamental level, the backreaction of matter on geometry cannot be neglected. Namely, geometry couples to matter through *Einstein's equations*

$$\boxed{\boldsymbol{R_{\mu\nu} - \tfrac{1}{2}R \cdot g_{\mu\nu} = \kappa T_{\mu\nu}[g]}}$$

and since matter undelies the rules of quantum mechanics, the right hand side of this equation, the stress-energy tensor $T_{\mu\nu}[g]$, becomes an operator. One has tried to keep geometry classical while matter is quantum mechanical by replacing $T_{\mu\nu}[g]$ by the Minkowski vacuum expectation value $< \hat{T}_{\mu\nu}[\eta] >$ but the solution of this equation will give $g \neq \eta$ which one then has to feed back into the definition of the vacuum expectation value etc. The resulting iteration does not converge in general. Thus, such a procedure is also inconsistent whence we *must quantize the gravitational field as well.* This leads to the *Quantum Einstein Equations*

$$\boxed{\boldsymbol{\hat{R}_{\mu\nu} - \tfrac{1}{2}\hat{R} \cdot \hat{g}_{\mu\nu} = \kappa \hat{T}_{\mu\nu}[\hat{g}]}}$$

Of course, this equation is only formal at this point and must be embedded in an appropriate Hilbert space context.

ii) *Inherent Classical Geometry Inconsistency*
Even without quantum theory at all Einstein's field equations predict spacetime singularities (black holes, big bang singularities etc.) at which the equations become meaningless. In a truly fundamental theory, there is no room for such breakdowns and it is suspected by many that the theory cures itself upon quantization in analogy to the Hydrogenium atom whose stability is classically a miracle (the electron should fall into the nucleus after a finite time lapse due to emission of Bremsstrahlung) but is easily explained by quantum theory.

iii) *Inherent Quantum Matter Inconsistency*
As outlined above, perturbative quantum field theory on curved spacetimes is itself also ill-defined due to its UV (short distance) singularities which can be cured only with an ad hoc recipe order by order which lacks a fundamental explanation, moreover, the perturbation series is presumably divergent. Besides that, the usually infinite vacuum energies being usually neglected in such a procedure contribute to the cosmological constant and should have a large gravitational backreaction effect. That such energy subtractions are quite significant is maybe best demonstrated by the Casimir effect. Now, since general relativity possesses a fundamental length scale, the Planck length, it has been argued ever since that gravitation plus matter should give a finite quantum theory since gravitation provides the necessary, built-in, short distance cut-off.

iv) *Perturbative Quantum Geometry Inconsistency*
Given the fact that perturbation theory works reasonably well if the coupling constant is small

for the non-gravitational interactions on a background metric it is natural to try whether the methods of quantum field theory on curved spacetime work as well for the gravitational field. Roughly, the procedure is to write the dynamical metric tensor as $g = \eta + h$ where η is the Minkowski metric and h is the deviation of g from it. One arrives at a formal, infinite series with finite radius of convergence which becomes meaningless if the fluctuations are large. Although the naive power counting argument implies that general relativity so defined is a non-renormalizable theory it was hoped that due to cancellations of divergencies the perturbation theory could be actually finite. However, that this hope was unjustified was shown in [9] where calculations demonstrated the appearance of divergencies at the two-loop-level, which suggests that at every order of perturbation theory one must introduce new coupling constants which the classical theory did not know about and one loses predictability.

It is well-known that the (locally) supersymmetric extension of a given non-supersymmetric field theory usually improves the ultra-violet convergence of the resulting theory as compared to the original one due to fermionic cancellations [10]. It was therefore natural to hope that quantized supergravity might be finite. However, in [11] a serious argument against the expected cancellation of perturbative divergences was raised and recently even the again popular (due to its M-Theory context) most complicated 11D "last hope" supergravity theory was shown not to have the magical cancellation property [12].

Summarizing, although a definite proof is still missing up to date (mainly due to the highly complicated algebraic structure of the Feynman rules for quantized supergravity) it is today widely believed that perturbative quantum field theory approaches to quantum gravity are meaningless.

The upshot of these considerations is that our understanding of quantum field theory and therefore fundamental physics is quite limited unless one quantizes the gravitational field as well. Being very sharply critical one could say:

The current situation in fundamental physics can be compared with the one at the end of the nineteenth century: While one had a successful theory of electromagnetism, one could not explain the stability of atoms. One did not need to worry about this from a practical point of view since atomic length scales could not be resolved at that time but from a fundamental point of view, Maxwell's theory was incomplete. The discovery of the mechanism for this stability, quantum mechanics, revolutionized not only physics. Still today we have no thorough understanding for the stability of nature (an experiment that everybody can repeat by looking out of the window) and it is similarly expected that the more complete theory of quantum gravity will radically change our view of the world. That is, considering the metric as a quantum operator will bring us beyond standard model physics even without the discovery of new forces, particles or extra dimensions.

The Role of Background Independence

The twentieth century has dramatically changed our understanding of nature : It revealed that physics is based on two profound principles : quantum mechanics and general relativity. Both principles revolutionize two pivotal structures of Newtonian physics : First, the determinism of Newton's equations of motion evaporates at a fundamental level, rather dynamics is reigned by probabilities underlying the Heisenberg uncertainty obstruction. Secondly, the notion of absolute time and space has to be corrected, space and time and distances between points of the spacetime manifold, that is, the metric, become themselves dynamical, geometry it is no longer just an observer. The usual Minkowski metric ceases to be a distinguished, externally prescribed, background structure. Rather, the laws of physics are *background independent*, mathematically expressed by the classical Einstein equations which are *generally (or four-diffeomorphism) covariant.* As we have argued, it is this new element of *background independence* brought in with Einstein's theory of gravity which completely changes our present understanding of quantum field theory.

A satisfactory physical theory must combine both of these fundamental principles, quantum mechanics and general relativity, in a consistent way and will be called "Quantum Gravity". However, the quantization of the gravitational field has turned out to be one of the most challenging unsolved problems in theoretical and mathematical physics. Although numerous proposals towards a quantization have been made since the birth of general relativity and quantum theory, none of them can be called successful so far. This is in sharp contrast to what we see with respect to the other three interactions whose description has culminated in the so-called standard model of matter, in particular, the spectacular success of perturbative quantum electrodynamics whose theoretical predictions could be verified *to all digits within the experimental error bars* until today.

Today we do not have a theory of quantum gravity, what we have is :
1) The Standard Model, a quantum theory of the non-gravitational interactions (electromagnetic, weak and strong) or *matter* which, however, completely ignores general relativity.
2) Classical General Relativity or *geometry*, which is a background independent theory of all interactions but completely ignores quantum mechanics.

What is so special about the gravitational force that it persists its quantization for about seventy years already ? As outlined in the previous subsection, *the answer is simply that today we only know how to do Quantum Field Theory (QFT) on fixed background metrics.* The whole formalism of ordinary QFT heavily relies on this background structure and collapses to nothing when it is missing. It is already much more difficult to formulate a QFT on a non-Minkowski (curved) background but it seems to become a completely hopeless task when the metric is a dynamical, even fluctuating quantum field itself. This underlines once more the source of our current problem of quantizing gravity : *We have to learn how to do QFT on a differential manifold (or something even more rudimentary, not even relying on a fixed topological, differentiable manifold) rather than a spacetime.*

In order to proceed, today a high energy physicist has the choice between the following two, extreme approaches :
Either the *particle physicist's*, who prefers to take over the well-established mathematical machinery from QFT on a background at the price of dropping background independence altogether to begin with and then tries to find the true background independent theory by summing the perturbation series (summing over all possible backgrounds). Or the *quantum geometer's*, who believes that background independence lies at the heart of the solution to the problem and pays the price to have to invent mathematical tools that go beyond the framework of ordinary QFT right from the beginning. Both approaches try to unravel the truly deep features that are unique to Einstein's theory associated with background independence from different ends.

The particle physicist's language is perturbation theory, that is, one writes the quantum metric operator as a sum consisting of a background piece and a perturbation piece around it, the graviton, thus obtaining a graviton QFT on a Minkowski background. *We see that perturbation theory, by its very definition, breaks background independence and diffeomorphism invariance at every finite order of perturbation theory.* Thus one can restore background independence only by summing up the entire perturbation series which is of course not easy. Not surprisingly, as already mentioned, applying this programme to Einstein's theory itself results in a mathematical desaster, a so-called non-renormalizable theory without any predictive power. In order to employ perturbation theory, it seems that one has to go to string theory which, however, requires the introduction of new additional structures that Einstein's classical theory did not know about: supersymmetry, extra dimensions and an infinite tower of new and very heavy particles next to the graviton. This is a fascinating but extremely drastic modification of general relativity and one must be careful not to be in conflict with phenomenology as superparticles, Kaluza Klein modes from the dimensional reduction and those heavy particles have not been observed until today. On the other hand, string theory has a good chance to be a unified theory of the perturbative aspects of all interactions in the sense that all interactions follow from a common object, the string, thereby explaining the particle content of the world.

The quantum geometer's language is a non-perturbative one, keeping background independence as a guiding principle at every stage of the construction of the theory, resulting in mathematical structures drastically different form the ones of ordinary QFT on a background metric. One takes Einstein's theory absolutely seriously, uses only the principles of general relativity and quantum mechanics and lets the theory build itself, driven by mathematical consistency. Any theory meeting these standards will be called *Quantum General Relativity (QGR)*. Since QGR does not modify the matter content of the known interactions, QGR is therefore not in conflict with phenomenology but also it cannot explain the particle content so far. However, it tries to unify all interactions in a different sense: all interactions must transform under a common gauge group, the four-dimensional diffeomorphism group which on the other hand is mlmost completely broken in perturbative approaches.

Let us remark that even without specifying further details, any QGR theory is a promising candidate for a theory that is free from two divergences of the so-called perturbation series of Feynman diagrammes common to all perturbative QFT's on a background metric: (1) Each term in the series diverges due to the ultraviolet (UV) divergences of the theory which one can cure for renormalizable theories, such as string theory, through so-called renormalization techniques and (2) the series of these renormalized, finite terms diverges, one says the theory is not finite. The first, UV, problem has a chance to be absent in a background independent theory for a simple but profound reason: In order to to say that a momentum becomes large one must refer to a background metric with respect to which it is measured, but there simply is no background metric in the theory. The second, convergence, problem of the series might be void as well since there are simply no Feynman diagrammes ! Thus, the mere existence of a consistent background independent quantum gravity theory could imply a finite quantum theory of all interactions.

Approaches to Quantum Gravity

The aim of the previous subsection was to convince the reader that background independence is, maybe, **The Key Feature** of quantum gravity to be dealt with. No matter how one deals with this issue, whether one starts from a perturbative (= background dependent) or from a non-perturbative (= background independent) platform, one has to invent something drastically new in order to quantize the gravitational field. We will now explain these approaches in more detail, listed in decreasing numbers of researchers working in the respective fields.

1) *Perturbative Approach : String Theory*
The only known consistent perturbative approach to quantum gravity is string theory which has good chances to be a theory that unifies all interactions. String Theory [13] is not a field theory in the ordinary sense of the word. Originally, it was a two-dimensional field theory of world-sheets embedded into a fixed, D-dimensional pseudo-Riemannian manifold (M, g) of Lorentzian signature which is to be thought of as the spacetime of the physical world. The Lagrangean of the theory is a kind of non-linear $\sigma-$model Lagrangean for the associated embedding variables X (and their supersymmetric partners in case of the superstring). If one perturbes $g(X) = \eta + h(X)$ as above and keeps only the lowest order in X one obtains a free field theory in two dimensions which, however, is consistent only when $D = 26$ (bosonic string) or $D = 10$ (superstring) respectively. Strings propagating in those dimensions are called critical strings, non-critiical strings exist but have so far not played a significant role due to phenomenological reasons. Remarkably, the mass spectrum of the particle-like excitations of the closed worldsheet theory contain a massless spin-two particle which one interprets as the graviton. Until recently, the superstring was favoured since only there it was believed to be possible to get rid off an unstable tachyonic vaccum state by the GSO projection. However, one recently also tries to construct stable bosonic string theories [14].

Moreover, if one incorporates the higher order terms $h(X)$ of the string action, sufficient for one loop corrections, into the associated path integral one finds a consistent quantum theory up to one loop only if the background metric satisfies the Einstein equations. These are the most powerful outcomes of the theory : although one started out with a fixed background metric, the background is not arbitrary but has to satisfy the Einstein equations up to higher loop corrections indicating that the one-loop effective action for the low energy quantum field theory in those D dimensions is Einstein's theory plus corrections. Finally, at least the type II superstring theories are are one-loop and, possibly, to all orders, *finite*. String theorists therefore argue to have found candidates for a consistent theory of quantum gravity with the additional advantage that they do not contain any free parameters (like those of the standard model) except for the string tension.

These facts are very impressive, however, some cautionary remarks are appropriate:

– *Vacuum Degeneracy*
Dimension $D+1 = 10, 26$ is not the dimension of everyday physics so that one has to argue that the extra $D - 3$ dimensions are "tiny" in the Kaluza-Klein sense although nobody knows the mechanism responsible for this "spontaneous compactification". According to [15] there exists an order of 10^4 consistent, distinct Calabi-Yau compactifications (other compactifications such as toroidal ones seem to be inconsistent with phenomenology) each of which has an order of 10^2 free, continuous parameters (moduli) like the vacuum expectation value of the Higgs field in the standard model. For each compactification of each of the five string theories in $D = 10$ dimensions and for each choice of the moduli one obtains a distinct low energy effective theory. This is clearly not what one expects from a theory that aims to unify all the interactions, the 18 (or more for massive neutrinos) free, continuous parameters of the standard model have been replaced by 10^2 continuous plus at least 10^4 discrete ones.

This vacuum degeneracy problem is not cured by the M-Theory interpretation of string theory but it is conceptually simplified if certain conjectures are indeed correct : String theorists believe (bearing on an impressively huge number of successful checks) that so-called T (or target space) and S (or strong – weak coupling) duality transformations between all these string theories exist which suggest that we do not have 10^4 unrelated 10^2-dimensional moduli spaces but that rather these 10^2-dimensional manifolds intersect in singular, lower dimensional submanifolds corresponding to certain singular moduli configurations. This typically happens when certain masses vanish or certain couplings diverge or vanish (in string theory the coupling is related to the vacuum expectation value of the dilaton field). Crucial in this picture are so-called D-branes, higher dimensional objects additional to strings which behave like solitons ("magnetic monopoles") in the electric description of a string theory and like fundamental objects ("electric degrees of freedom") in the S-Dual description of the same string theory, much like the electric – magnetic duality of Maxwell theory under which strong and weak coupling are exchanged. Further relations between different string theories are obtained by compactifying them in one way and decompactifying them in another way, called a T – duality transformation. The resulting picture is that there exists only one theory which has all these compactification limits just described, called M – Theory. Curiously, M – Theory is an 11D theory whose low energy limit is 11D supergravity and whose weak coupling limit is type IIA superstring theory (obtained by one of these singular limits since the size of the 11th compactified dimension is related to the string coupling again). Since 11D supergravity is also the low energy limit of the 11D supermembrane, some string theorists interpret M-Theory as the quantized 11D supermembrane (see, e.g., [16] and references therein).

– *Phenomenology Match*
  Until today, no conclusive proof exists that for any of the compactifications described above we obtain a low energy effective theory which is experimentally consistent with the data that we have for the standard model [17] although one seems to get at least rather close. The challenge in string phenomenology is to consistently and spontaneously break supersymmetry in order to get rid off the so far non-observed superpartners. There is also an infinite tower of very massive (of the order of the Planck mass and higher) excitations of the string but these are too heavy to be observable. More interesting are the Kaluza Klein modes whose masses are inverse proportional to the compactification radii and which have recently given rise to speculations about "sub-mm-range" gravitational forces [18] which one must make consistent with observation also.

– *Fundamental Description*
  Even before the M – Theory revolution, string theory has always been a theory without Lagrangean description, S – Matrix element computations have been guided by conformal invariance but there is no "interaction Hamiltonian", string theory is a first quantized theory. Second quantization of string theory, called string field theory [19], has so far not attracted as much attention as it possibly deserves. However, a fascinating possibility is that the 11D supermembrane, and thus M – Theory, is an already second quantized theory [20].

– *Background Dependence*
  As mentioned above, string theory is best understood as a free 2D field theory propagating on a 10D Minkowski target space plus perturbative corrections for scattering matrix computations. This is a heavily background dependent description, issues like the action of the 10D diffeomorphism group, the fundamental symmetry of Einstein's action, or the

backreaction of matter on geometry, cannot be asked. Perturbative string theory, as far as quantum gravity is concerned, can describe graviton scattering in a background spacetime, however, most problems require a non-perturbative description when the backreaction can no longer be ignored, such as scattering at quantum black holes. As a first step in that direction, recently stringy black holes have been discussed [21]. Here one uses so-called BPS D-brane configurations which are so special that one can do a perturbative calculation and extend it to the non-perturbative regime since the results are protected against non-perturbative corrections due to supersymmetry. So far this works only for extremely charged, supersymmetric black holes which are astrophysically not very realistic. But still these developments are certainly a move in the right direction since they use for the first time non-perturbative ideas in a crucial way and have been celebrated as one of the triumphes of string theory.

2) *Non-Perturbative Approaches*
The non-perturbative approaches to quantum gravity can be grouped into the following five main categories.

2a) *Canonical Quantum General Relativity*
If one wanted to give a definition of this theory then one could say the following:

Canonical Quantum General Relativity is an attempt to construct a mathematically rigorous, non-perturbative, background independent quantum theory of four-dimensional, Lorentzian general relativity plus all known matter in the continuum.

This is the oldest approach and goes back to the pioneering work by Dirac [22] started in the 40's and was further developed especially by Wheeler and DeWitt [23] in the 60's. The idea of this approach is to apply the Legendre transform to the Einstein-Hilbert action by splitting spacetime into space and time and to cast it into Hamiltonian form. The resulting "Hamiltonian" H is actually a so-called Hamiltonian constraint, that is, a Hamiltonian density which is constrained to vanish by the equations of motion. A Hamiltonian constraint must occur in any theory that, like general relativity, is invariant under local reparameterizations of time. According to Dirac's theory of the quantization of constrained Hamiltonian systems, one is now supposed to impose the vanishing of the quantization $\hat{H}$ of the Hamiltonian constraint H as a condition on states ψ in a suitable Hilbert space $\mathcal{H}$, that is, formally

$$\boxed{\hat{H}\psi = 0}$$

This is the famous Wheeler-DeWitt equation or **Quantum-Einstein-Equation** of canonical quantum gravity and resembles a Schrödinger equation, only that the familiar $\partial\psi/\partial t$ term is missing, one of several occurences of the "absence or problem of time" in this approach (see, e.g., [29] and references therein).
Since the status of this programme is the subject of the present review we will not go too much into details here. The successes of the theory are a mathematically rigorous framework, manifest background independence, a manifestly non-perturbative language, an inherent notion of quantum discreteness of spacetime which is *derived* rather than pos-

tulated, certain UV finiteness results, a promising path integral formulation (spin foams) and finally a consistent formulation of quantum black hole physics.

The following issues are at the moment unresolved within this approach:

* *Tremendously Nonlinear Structure*
  The Wheeler-DeWitt operator is, in the so-called ADM formulation, a functional differential operator of second order of the worst kind, namely with non-polynomial, not even analytic (in the basic configuration variables) coefficients. To even define such an operator rigorously has been a major problem for more than 60 years. What should be a suitable Hilbert space that carries such an operator ? It is known that a Fock Hilbert space is not able to support it. Moreover, the structure of the solution space is expectedly very complicated. Thus we see that one meets a great deal of mathematical problems before one can even start addressing physical questions. As we will describe in this report there has been a huge amount of progress in this direction since the introduction of new canonical variables due to Ashtekar [30] in 1986. However, the physics of the Wheeler-DeWitt operator is still only poorly understood.
* *Loss of Manifest Four-Dimensional Diffeomorphism Covariance*
  Due to the split of spacetime into space and time the treatment of spatial and time diffeomorphisms is somewhat different and the original four-dimensional covariance of the theory is no longer manifest. Classically one can prove (and we will in fact do that later on) that four-dimensional diffeomorphism covariance is encoded in a precise sense into the canonical formalism, although it is deeply hidden. In quantum theory the proper implementation of the diffeomorphism group is the question whether the so-called Dirac algebra (of which the Wheeler – DeWitt operator is an element) has an anomaly or not and which at the moment has no conclusive answer.
  Let us clarify an issue that comes up often in debates between quantum geometers and string theorists:
  What one means by $(D+1)$-dimensional covariance in string theory on a Minkowski target space is just $(D+1)-$dimensional *Poincaré* covariance but not *Diffeomorphism* covariance. Clearly the Poincaré group is a subgroup of the diffeomorphism group (for asymptotically flat spacetimes) of measure zero and the rest of the diffeomorphism group, which is in fact the symmetry group of Einstein's theory, is completely broken in string theory. In canonical QGR at least the huge spatial subgroup of the diffeomorphism group is manifestly without anomalies and possibly the remaining part of the diffeomorphism group as well.
* *Interpretational (Conceptual) Issues*
  Once one has found the solutions of the Quantum Einstein Equations one must find a complete set of Dirac observables (operators that leave the space of solutions invariant) which is an impossible task to achieve even in classical general relativity. One must therefore find suitable approximation methods which is a development which has just recently started. However, even if one would have found those (approximate) operators, which would be in some sense even time independent and therefore extremely non-local, one would need to deparameterize the theory, that is, one must find an explanation for the local dynamics in our world. There are some proposals for dealing with this issue but there is no rigorous framework available at the moment.
* *Classical Limit*
  As we will see, our Hilbert space is of a new (background independent) kind, operators

are regulated in a non-standard (background independent) way. It is therefore no longer clear that the theory that has been constructed so far indeed has general relativity as its classical limit. Again, semiclassical analysis has just been launched recently.

2b) *Continuum Functional Integral Approach*
Here one tries to give meaning to the sum over histories of e^{-S_E} where S_E denotes the Euclidean Einstein-Hilbert action [31]. It is extremely hard to do the path integral and apart from semi-classical approximations and steepest descent methods in simplified models with a finite number of degrees of freedom one could not get very far within this framework yet. There are at least the two following reasons for this :
1) The action functional S_E is unbounded from below. Therefore the path integral is badly divergent from the outset and although rather sophisticated proposals have been made of how to improve the convergence properties, none of them has been fully successful to the best knowledge of the author.
2) The Euclidean field theory underlying the functional integral and the quantum theory of fields propagating on a Minkowski background are related by Wick rotating the Schwinger functions of the former into the Wightman functions of the latter (see, e.g., [32]). However, in the case of quantum gravity the metric itself becomes dynamical and is being integrated over, therefore the concept of Wick rotation becomes ill-defined. In other words, there is no guarantee that the Euclidean path integral even has any relevance for the quantum field theory underlying the Lorentzian Einstein-Hilbert action.
On the other hand, the functional integral approach has motivated the consistent histories approach to the quantum mechanics of closed systems (cosmologies) [33] which in many senses is superiour over the Copenhagen interpretation.

2c) *Lattice Quantum Gravity*
This approach can be subdivided into two main streams (see [34] for a review):
a) Regge Calculus [35]. Here one introduces a fixed triangulation of spacetime and integrates with a certain measure over the lengths of the links of this triangulation. The continuum limit is reached by refining the triangulation.
b) Dynamical Triangulations [36]. Here one takes the opposite point of view and keeps the lengths of the links fixed but sums over all triangulations. The continuum limit is reached by taking the link length to zero.
In both approaches one has to look for critical points (second order phase transitions). An issue in both approaches is the choice of the correct measure. Although there is no guideline, it is widely believed that the dependence on the measure is weak due to universality in the statistical mechanical sense. The reason for the possibility that the path integral exists although the Euclidean action is unbounded from below is that the configurations with large negative action have low volume (measure) so that "entropy wins over energy". Especially in the field of dynamical triangulations there has been a major breakthrough recently [37]: The convergence of the partition function could be established *analytically* in two dimensions (the action is basically a cosmological constant term) and the relation between the Lorentzian and Euclidean theory becomes transparent. This opens the possibility that similar results hold in higher dimensions, in particular, it seems as if the Lorentzian theory is much better behaved than the Euclidean theory because one has to sum over fewer configurations (those that are compatible with quantum causality). There are also promising new results concerning a non-perturbative Wick rotation [38].

What is still missing within this approach (in more than two dimensions), as in any path integral approach for quantum gravity that has been established so far, is a clear physical interpretation of the expectation values of observables as transition amplitudes in a given Hilbert space. A possible way out could be proposed if one were able to establish reflection positivity of the measure, see [32].

2d) *Non-Orthodox Approaches*
Approaches belonging to this approach start by questioning standard quantum field theory at an even more elementary level. Namely, if the ideas about spacetime foam (discrete structure of spacetime) are indeed true then *one should not even start formulating quantum field theory on a differentiable manifold* but rather something intrinsically discrete. Maybe we even have to question the foundations of quantum mechanics and to depart from a purely binary logic. To this category belong the *Non-Commutative Geometry* by Alain Connes [39] also considered recently by string theorists [40], the *Topos Theory* by Chris Isham [41], the *Twistor Theory* by Roger Penrose [42], the *Causal Set Programme* by Raphael Sorkin [43] and finally the Consistent History approach due to Gell-Mann and Hartle [33] which recently has picked up some new momentum in terms of the history phase space due to Isham et al and Kuchař et al [44].
These approaches are, maybe, the most radical reformulations of fundamental physics but they are also the most difficult ones because the contact with standard quantum field theory is very small. Consequently, these programmes are in some sense "farthest" from observation and are consequently least developed so far. However, the ideas spelled out in these programmes could well reappear in the former approaches as well once the latter have reached a sufficiently high degree of maturity.

All of these four nonperturbative programmes are mutually loosely connected : Roughly, the operator formulation of the canonical approach is equivalent to the continuous path integral formulation through some kind of Feynman-Kac formula a concrete implementation of which are the so-called spin-foam models to be mentioned later, lattice quantum gravity is a discretization of the path integral formulation and both the canonical and the lattice approach seem to hint at discrete structures on which the non-orthodox programmes are based.

Finally, every non-perturbative programme better contains a sector which is well described by perturbation theory and therefore string theory which then provides an interface between the two big research streams.

This ends our survey of the existent quantum gravity programmes. By far most of the people, of the order of 10^3, work in string theory, followed by the canonical programme with an order of 10^2 scientists, then the lattice quantum gravity –, functional integral – and the non-commutative programme with an order of 10^1 researchers and finally the non-orthodox programmes except for the non-commutative one with an order of 10^0 physicists.

Motivation for Canonical Quantum General Relativity

We close this section by motivating our choice to follow the non-perturbative, canonical approach. Of course, our discussion cannot be entirely objective.

I) *Non-Perturbative versus Perturbative*

Our preference for a nonperturbative approach is twofold:

The first reason is certainly a matter of taste, a preference for a certain methodology: Try to combine the two fundamental principles, general relativity and quantum mechganics *with no additional structure*, explore the logical consequences and push the framework until success or until there is a contradiction (inconsistency) either within the theory or with the experiment. In the latter case, examine the reason for failure and try to modify theory appropriately. The reason for not allowing additional structure (principle of minimality) is that unless we only use structures which have been confirmed to be a property of nature then we are standing in front of an ocean of possible new theories which a priori could be equally relevant. In a sense we are saying that if gravity cannot be quantized perturbatively without extra structures such as necessary in string theory, then one should try a nonperturbative approach. If that still fails then maybe we find out why and exactly which extra structures are necessary rather than guessing them. Such a methodology has proved to be very successful in the history of science.

The second reason, however, is maybe more serious: *It is not at all true that perturbation theory is always a good approximation in a non-empty neighbourhood of the expansion point.* To quote an example from [3], consider the harmonic oscillator Hamiltonian $H = p^2 + \omega^2 q^2$ and let us treat the potential $V = \omega^2 q^2$ as an interaction Hamiltonian perturbing the free Hamiltonian $H_0 = p^2$ at least for low frequencies ω. The exact spectrum of H is discrete while that of H_0 is continuous. The point is now that *one is never going to see, for no value of the "coupling constant" $\omega > 0$, the discreteness of the unperturbed Hamiltonian by doing perturbation theory and thus one completely misses the correct physics* !

Finally, borrowing from [45], let us exhibit a calculation which demonstrates the *regularizing mechanism of a non-perturbative treatment of general gravity taking its very non-linear nature very serious.*
Consider the self-energy of a bare point charge e_0 with rest mass m_0 due to static electromagnetic and gravitational interaction. From the point of view of Newtonian physics, this energy is of the form ($\hbar = c = 1$, the bare Newton's constant is denoted by G_0)

$$m(r) = m_0 + e_0^2/r - G_0 m_0^2/r$$

and diverges as $r \to 0$ unless e_0, m_0, G_0 are fine tuned. However, general relativity tells us that all of the mass of the charge, that is rest mass plus field energy within a shell of radius r couples to the gravitational field which is why above equation should be replaced by

$$m(r) = m_0 + e_0^2/r - G_0 m(r)^2/r$$

which can be solved for

$$m(r) = \frac{r}{2G_0}[-1 + \sqrt{1 + \frac{4G_0}{r}(m_0 + \frac{e_0^2}{r})}]$$

Notice that now the bare mass $m(r=0) = e_0/\sqrt{G_0}$ is *finite without fine tuning. Moreover, the result is non-analytical in Newton's constant G_0 and is not accessible by perturbation theory, in particular, the bare mass is independent of the rest mass* ! Of course, this calculation should not be taken too seriously since e.g. no quantum effects have been brought in, it merely serves to illustrate our point that general relativity could serve as natural regulator of field theory divergences. (However, a proper general relativistic treatment (ADM mass of the Reissner-Nordstrøm solution) [45] can be performed, see also for more details).

These arguments can be summarized by saying that there is a good chance that perturbative quantum gravity *completely misses the point* although, of course, there is no proof !

II) *Canonical versus other Non-Perturbative Approaches*

Here our motivation is definitely just a matter of taste, that is, we take a practical viewpoint:

Path integrals have the advantage that they are manifestly four-dimensionally diffeomorphism invariant but their huge disadvantage is that they are awfully hard to compute analytically, even in quantum mechanics. While numerical methods will certainly enter the canonical approach as well in the close future one gets farther with analytical methods. However, it should be stressed that path integrals and canonical methods are very closely related and usually one can derive one from the other through some kind of Feynman – Kac formula.

The non-orthodox approaches have the advantage of starting from a discrete/non-commutative spacetime structure from scratch while in canonical quantum gravity one begins with a smooth spacetime manifold and obtains discrete structures as a derived concept only which is logically less clean: The true theory is the quantum theory and if the world is discrete one should not begin with smooth structures at all. Our viewpoint is here that, besides the fact that again the canonical approach is more minimalistic, at some stage in the development of the theory there must be a quantum leap and in the final reformulation of the theory everything is just combinatorical. This can actually be done in 2+1 gravity as we will describe later on !

Notation and Conventions

You can tell whether a high energy physicist is a particle physicist or a quantum geometer from the index notation that she or he uses. We obviously use here the quantum geometer's (that is, latin letters from the beginning of the alphabet are tensorial while those from the middle are Lie algebra indices), the particle physicists's is often just opposite.

G$= 6.67 \cdot 10^{-11} m^3 kg^{-1} s^{-2}$: Newton's constant
$\kappa = 8\pi \mathrm{G}/c^3$: Gravitational coupling constant
$\ell_p = \sqrt{\hbar\kappa} = 10^{-33}$ cm: Planck length
$m_p = \sqrt{\hbar/\kappa}/c = 10^{19}$GeV: Planck mass
Q: Yang-Mills coupling constant
$M,\ \dim(M) = D+1$: Spacetime manifold
$\sigma,\ \dim(\sigma) = D$: Abstract spatial manifold
Σ: Spatial manifold embedded into M
G: Compact gauge group
Lie(G): Lie algebra
$N-1$: Rank of gauge group
$\mu,\nu,\rho,.. = 0,1,..,D$: Tensorial spacetime indices
$a,b,c,.. = 1,..,D$: Tensorial spatial indices
$\epsilon_{a_1..a_D}$: Levi-Civita totally skew tensor density
$g_{\mu\nu}$: Spacetime metric tensor
q_{ab}: Spatial (intrinsic) metric tensor of σ
K_{ab}: Extrinsic curvature of σ
R: Curvature tensor
$\underline{h}$: Group elements for general G
$\underline{h}_{mn},\ m,n,o,.. = 1,..,N$: Matrix elements for general G
$I,J,K,.. = 1,2,..,\dim(G)$: Lie algebra indices for general G
$\underline{\tau}_I$: Lie algebra generators for general G
$k_{IJ} = -\mathrm{tr}(\underline{\tau}_I\underline{\tau}_J)/N := \delta_{IJ}$: Cartan-Killing metric for G
$[\underline{\tau}_I,\underline{\tau}_J] = 2f_{IJ}\ ^K\underline{\tau}_K$: Structure constants for G
$\underline{\pi}(\underline{h})$: (Irreducible) representations for general G
h: Group elements for $SU(2)$
$h_{AB},\ A,B,C,.. = 1,2$: Matrix elements for $SU(2)$
$i,j,k,.. = 1,2,3$: Lie algebra indices for $SU(2)$
τ_i: Lie algebra generators for $su(2)$
$k_{ij} = \delta_{ij}$: Cartan-Killing metric for $SU(2)$
$f_{ij}\ ^k = \epsilon_{ijk}$: Structure constants for $SU(2)$
$\pi_j(h)$: (Irreducible) representations for $SU(2)$ with spin j
$\underline{A}$: Connection on G-bundle over σ
$\underline{A}_a^I$: Pull-back of $\underline{A}$ to σ by local section
g: gauge transformation or element of complexification of G
P: Principal $G-$bundle
A: Connection on $SU(2)$-bundle over σ
A_a^i: Pull-back of A to σ by local section
$*\underline{E}$: $(D-1)$-form covector bundle associated to the G-bundle under the adjoint representation
$*\underline{E}^I_{a_1..,a_{D-1}} =: k^{IJ}\epsilon_{a_1..,a_D}\underline{E}^{a_D}_J$: Pull-back of $*\underline{E}$ to σ by local section

$*E$: $(D-1)$-form covector bundle associated to the $SU(2)$-bundle under the adjoint representation
$*E^i_{a_1..,a_{D-1}} =: k^{ij}\epsilon_{a_1..,a_D} E^{a_D}_j$: Pull-back of $*E$ to σ by local section
$E^a_j := \epsilon^{a_1..a_{D-1}}(*E)^k_{a_1..a_{D-1}} k_{jk}/((D-1)!)$: "Electric fields"
e: One-form covector bundle associated to the $SU(2)$-bundle under the defining representation (D-bein)
e^i_a: Pull-back of e to σ by local section
Γ^i_a: Pull-back by local section of $SU(2)$ spin connection over σ
$\mathcal{M}$: Phase space
$\mathcal{E}$: Banach manifold or space of smooth electric fields
$T_{(a_1..a_n)} := \frac{1}{n!}\sum_{\iota\in S_n} T_{a_{\iota(1)}..a_{\iota(n)}}$: Symmetrization of indices
$T_{[a_1..a_n]} := \frac{1}{n!}\sum_{\iota\in S_n} \mathrm{sgn}(\iota)\, T_{a_{\iota(1)}..a_{\iota(n)}}$: Antisymmetrization of indices
$\mathcal{A}$: Space of smooth connections or abstract algebra
$\mathcal{G}$: Space of smooth gauge transformations
$\overline{\mathcal{A}}$: Space of distributional connections
$\mathcal{A}^{\mathbb{C}}$: Space of smooth complex connections
$\mathcal{G}^{\mathbb{C}}$: Space of smooth complex gauge transformations
$\overline{\mathcal{G}}$: Space of distributional gauge transformations
$\mathcal{A}/\mathcal{G}$: Space of smooth connections modulo smooth gauge transformations
$\overline{\mathcal{A}}/\overline{\mathcal{G}}$: Space of distributional connections modulo distributional gauge transformations
$\overline{\mathcal{A}/\mathcal{G}}$: Space of distributional gauge equivalence classes of connections
$\overline{\mathcal{A}}^{\mathbb{C}}$: Space of distributional complex connections
$\overline{\mathcal{A}/\mathcal{G}}^{\mathbb{C}}$: Space of distributional complex gauge equivalence classes of connections
$\mathcal{C}$: Set of piecewise analytic curves or classical configuration space
$\overline{\mathcal{C}}$: Quantum configuration space
$\mathcal{P}$: Set of piecewise analytic paths
$\mathcal{Q}$: Set of piecewise analytic closed and basepointed paths
$\mathcal{L}$: Set of tame subgroupoids of $\mathcal{P}$ or general label set
$\mathcal{S}$: Set of tame subgroups of $\mathcal{Q}$ (hoop group) or set of spin network labels
s: spin-net= spin-network label (spin-net)
Γ^ω_0: Set of piecewise analytic, compactly supported graphs
Γ^ω_σ: Set of piecewise analytic, countably infinite graphs
c: Piecewise analytic curve
p: Piecewise analytic path
e: Entire analytic path (edge)
α: Entire analytic closed path (hoop)
γ: Piecewise analytic graph
v: vertex of a graph
$E(\gamma)$: Set of edges of γ
$V(\gamma)$: Set of vertices of γ
$h_p(A) = A(p)$: holonomy of A along p
$\prec$: abstract partial order
Ω: vector state or symplectic structure or curvature two-form
F: pull-back to σ of 2Ω by a local section
ω: general state
$\mathcal{H}$: general Hilbert space
Cyl: Space of cylindrical functions

$\mathcal{D}$: dense subspace of $\mathcal{H}$ equipped with a stronger topology
$\mathcal{D}'$: topological dual of $\mathcal{D}$
$\mathcal{D}^*$: algebraic dual of $\mathcal{D}$
$\mathcal{H}^0 = L_2(\overline{\mathcal{A}}, d\mu_0)$: Uniform measure L_2 space
$\mathcal{H}^{\otimes}$: Infinite Tensor Product extension of $\mathcal{H}^0$
Cyl_l: Restriction of Cyl to functions cylindrical over γ
$[.], (.)$: Equivalence classes
$\mathrm{Diff}^{\omega}(\sigma)$: Group of analytic diffeomorphisms of σ
φ: analytic diffeomorphism

DISCLAIMER

Like every review also this one is necessarily incomplete. There are many more fascinating results which we could not possibly also describe due to lack of space and time.

The selection of material is certainly biased by the author's own taste and by the fact that we wanted to keep this review at the same time compact and complete. This is not meant, at all, as a quality evaluation. We apologize to all those researchers whose results were, in their mind, not (or not sufficiently) highlighted and hope that the extensive bibliography displays a reasonably fair and objective overview of the available literature.

In that respect, let us mention that until four years ago a complete and nicely structured literature list [280] was available which unfortunately has not been updated since then. It would be nice if some good soul would find the time and strength to fill in the huge amount of literature that has piled up over the last four years.

Part I

Foundations

I.1 Classical Hamiltonian Formulation of General Relativity and the Programme of Canonical Quantization

In this section we focus on the classical Hamiltonian formulation of general relativity. First we repeat the most important steps that have lead to the metrical formulation due to Arnowitt, Deser and Misner [46]. Then we introduce the main ideas behind the programme of canonical quantization and summarize the status of that programme when applied to general relativity in the ADM formulation. As we will see, not much progress could be achieved in that formulation which has motivated Ashtekar to look for a different one, more suitable to quantization. We introduce this connection formulation in the final subsection of this section.

I.1.1 The ADM Formulation

The major reference for this subsection is definitely the beautiful textebook by Wald (especially appendix E and chapter 10) and by Hawking&Ellis [47].

The object of interest is the Einstein – Hilbert action for metric tensor fields $g_{\mu\nu}$ of Lorentzian ($s=-1$) or Euclidean ($s=+1$) signature which propagate on a $(D+1)$-dimensional manifold M

$$S=\frac{1}{\kappa}\int_M d^{D+1}X\sqrt{|\det(g)|}R^{(D+1)}\,. \tag{I.1.1. 1}$$

In this article we will be mostly concerned with $s=-1, D=3$ but since the subsequent derivations can be done without extra effort we will be more general here. Our signature convention is "mostly plus, that is, $(-,+,..,+)$ or $(+,+,..,+)$ in the Lorentzian or Euclidean case respectively so that timelike vectors have negative norm in the Lorentzian case.. Here $\mu,\nu,\rho,..=0,1,..,D$ are indices for the components of spacetime tensors and X^μ are the coordinates of M in local trivializations. $R^{(D+1)}$ is the curvature scalar associated with $g_{\mu\nu}$ and $\kappa=8\pi\mathrm{G}$ where G is Newton's constant (in units where $c=1$). The definition of the Riemann curvature tensor is in terms of one-forms given by

$$[\nabla_\mu,\nabla_\nu]u_\rho={}^{(D+1)}R_{\mu\nu\rho}{}^{\sigma}\,u_\sigma \tag{I.1.1. 2}$$

where ∇ denotes the unique, torsion-free, metric-compatible, covariant differential associated with $g_{\mu\nu}$. To make the action principle corresponding to (I.1.1. 1) well-defined one has, in general, to add boundary terms which we avoid by assuming that M is spatially compact without boundary. The more general case can be treated similarly, see e.g. [48], but would unnecessarily complicate the analysis.

In order to cast (I.1.1. 1) into canonical form one makes the assumption that M has the special topology $M=\mathbb{R}\times\sigma$ where σ is a fixed three-dimensional, compact manifold without boundary. By a theorem due to Geroch [49], if the spacetime is globally hyperbolic (existence of Cauchy surfaces in accordance with the determinism of classical physics) then it is necessarily of this kind of topology. Therefore, for classical physics our assumptions about the topology of M seems to be no restriction at all, at least in the Lorentzian signature case. In quantum gravity, however, different kinds of topologies and, in particular, *topology changes* are conceivable. Our philosophy will be first to construct the quantum theory of the gravitational field based on the classical assumption that $M=\mathbb{R}\times\sigma$ and then to *lift this restriction* in the quantum theory. A concrete proposal for such a lifting which naturally suggests itself in our formulation will be given later on, we are even able to allow for certain classes of signature changes ! Notice that *none* of the approaches listed in the introduction, except for the path integral approach which, however, is mathematically poorly defined in more than two spacetime dimensions, knows how to deal with topology changes from first principles, there exist only ad hoc prescriptions.

Having made this assumption, one knows that M foliates into hypersurfaces $\Sigma_t := X_t(\sigma)$, that is, for each fixed $t \in \mathbb{R}$ we have an embedding (a globally injective immersion) $X_t : \sigma \to M$ defined by $X_t(x) := X(t,x)$ where x^a, $a,b,c,.. = 1,2,3$ are local coordinates of σ. Likewise we have a diffeomorphism $X : \mathbb{R}\times\sigma \mapsto M;\ (t,x) \mapsto X(t,x) := X_t(x)$, in other words, a one-parameter family of embeddings is equivalent to a diffeomorphism. We would like to use these special diffeomorphisms in order to give a $D+1$ (space and time) decomposition of the action (I.1.1. 1). Now, since the action (I.1.1. 1) is invariant under all diffeomorphisms (changes of the coordinate system) of M the families of embeddings X_t are not specified by it and we allow them to completely arbitrary (a precise characterization of these "embedding diffeomorphisms" as compared to Diff(M) can be found in [50]). A useful parameterization of the embedding and its arbitrariness can be given through its deformation vector field

$$T^\mu(X) := (\frac{\partial X^\mu(t,x)}{\partial t})_{|X=X(x,t)} =: N(X)n^\mu(X) + N^\mu(X) \qquad \text{(I.1.1. 3)}$$

Here n^μ is a unit normal vector to Σ_t, that is, $g_{\mu\nu}n^\mu n^\nu = s$ and N^μ is tangential, $,g_{\mu\nu}n^\mu X^\nu_{,a} = 0$. The coefficients of proportionality N and N^μ respectively are called *lapse* function and *shift* vector field respectively. Notice that implicitly information about the metric $g_{\mu\nu}$ has been invoked into (I.1.1. 3), namely we are only dealing with spacelike embeddings and metrics of the above specified signature. The lapse is nowhere vanishing since for a foliation T must be timelike everywhere. Moreover, we take it to be *positive* everywhere as we want a future directed foliation (negative sign would give a past directed one and mixed sign would not give a foliation at all since then necessarily the leaves of the foliation would intersect).

We need one more property of n : By the inverse function theorem, the surface Σ_t can be defined by an equation of the form $f(X) = t = const.$. Thus, $0 = \lim_{\epsilon\to 0}[f(X_t(x+\epsilon b) - f(X_t(x))]/\epsilon = b^a X^\mu_{,a}(f_{,\mu})_{X=X_t(x)}$ for any tangential vector b of σ in x. It follows that up to normalization the normal vector is proportional to an exact one-form, $n_\mu = F f_{,\mu}$ or, in the language of forms, $n = n_\mu dX^\mu = Fdf$. Actually, this fact is an easy corollary from Frobenius' theorem (the surfaces Σ_t are the integral manifolds of the distribution $v:\ M \mapsto T(M);\ X \mapsto V_X(n) = \{v \in T_X(M); i_v(n) = 0\} \subset T_X(M)$).

Let us forget about the foliation for a moment and just suppose that we are given a hypersurface σ embedded into M via the embedding X. Let n be its unit normal vector field and $\Sigma = X(\sigma)$ its image. We now have the choice to work either on σ or on Σ when developing the tensor calculus of so-called spatial tensor fields. To work on Σ has the advantage that we can compare spatial tensor fields with arbitrary tensor tensor fields *restricted to* Σ because they are both tensor fields on a subset of M. Moreover, once we have developed tensor calculus on Σ we immediately have the one on σ by just pulling back (covariant) tensor fields on Σ to σ via the embedding, see below.

Consider then the following tensor fields, called the *first and second fundamental form* of Σ

$$q_{\mu\nu} := g_{\mu\nu} - s n_\mu n_\nu \text{ and } K_{\mu\nu} := q^\rho_\mu q^\sigma_\nu \nabla_\rho n_\sigma \qquad \text{(I.1.1. 4)}$$

where all indices are moved with respect to $g_{\mu\nu}$. Notice that both tensors in I.1.1. 4, are "spatial", i.e. they vanish when either of their indices is contracted with n^μ. A crucial property of $K_{\mu\nu}$ is its symmetry : We have $K_{[\mu\nu]} = q^\rho_\mu q^\sigma_\nu((\nabla_{[\rho}\ln(F))n_{\sigma]} + F\nabla_{[\mu}\nabla\nu]f) = 0$ since ∇ is torsion free. The square brackets denote antisymmetrization defined as an idempotent operation. From this fact one derives another useful differential geometric identity by employing the relation between the covariant differential and the Lie derivative :

$$2K_{\mu\nu} = q^\rho_\mu q^\sigma_\nu(2\nabla_{(\rho}n_{\sigma)})$$

$$
\begin{aligned}
&= q_\mu^\rho q_\nu^\sigma (\mathcal{L}_n g)_{\rho\sigma} = q_\mu^\rho q_\nu^\sigma (\mathcal{L}_n q + s\mathcal{L}_n n \otimes n)_{\rho\sigma} \\
&= q_\mu^\rho q_\nu^\sigma (\mathcal{L}_n q)_{\rho\sigma} = (\mathcal{L}_n q)_{\mu\nu} \qquad \text{(I.1.1. 5)}
\end{aligned}
$$

since $n^\mu \mathcal{L}_n q_{\mu\nu} = -q_{\mu\nu}[n,n]^\mu = 0$. Using $n^\mu = (T^\mu - N^\mu)/N$ we can write (I.1.1. 5) in the form

$$
2K_{\mu\nu} = \frac{1}{N}(\mathcal{L}_{T-N} q)_{\mu\nu} - 2n^\rho q_{\rho(\mu} \ln(N)_{,\nu)} = \frac{1}{N}(\mathcal{L}_{T-N} q)_{\mu\nu} \qquad \text{(I.1.1. 6)}
$$

Next we would like to construct a covariant differential associated with the metric $q_{\mu\nu}$. We would like to stress that this metric is non-degenerate as a bijection between *spatial tensors only* and not as a metric between arbitrary tensors defined on Σ. Recall that, by definition, a differential ∇ is said to be covariant with respect to a metric g (of any signature) on a manifold M if it is 1) metric compatible, $\nabla g = 0$ and 2) torsion free, $[\nabla_\mu, \nabla_\nu] f = 0 \; \forall C^\infty(M)$. According to a classical theorem, these two conditions fix ∇ uniquely in terms of the Christoffel connection which in turn is defined by its action on one-forms through $\nabla_\mu u_\nu := \partial_\mu u_\nu - \Gamma^\rho_{\mu\nu} u_\rho$. Since the tensor q is a metric of Euclidean signature on Σ we can thus apply these two conditions to q and we are looking for a covariant differential D on spatial tensors only such that 1) $D_\mu q_{\nu\rho} = 0$ and 2) $D_{[\mu} D_{\nu]} f = 0$ for scalars f. Of course, the operator D should preserve the set of spatial tensor fields. It is easy to verify that $D_\mu f := q_\mu^\nu \nabla_\nu \tilde{f}$ and $D_\mu u_\nu := q_\mu^\rho q_\nu^\sigma \nabla_\rho \tilde{u}_\sigma$ for $u_\mu n^\nu = 0$ and extended to arbitrary tensors by linearity and Leibniz' rule, does the job and thus, by the above mentioned theorem, is the unique choice. Here, $\tilde{f}$ and $\tilde{u}$ denote arbitrary smooth extensions of f and u respectively into a neighbourhood of Σ in M, necessary in order to perform the ∇ operation. The covariant differential is independent of that extension as derivatives not tangential to Σ are projected out by the q tensor (go into a local, adapted system of coordinates to see this) and we will drop the tilde again. One can convince oneself that the action of D on arbitrary spatial tensors is then given by acting with ∇ in the usual way followed by spatial projection of all appearing indices including the one with respect to which the derivative was taken.

We now ask what the Riemann curvature $R^{(D)\ \sigma}_{\mu\nu\rho}$ of D is in terms of that of ∇. To answer this question we need the second covariant differential of a spatial co-vector u_ρ which when carefully using the definition of D is given by

$$
\begin{aligned}
D_\mu D_\nu u_\rho &= q_\mu^{\mu'} q_\nu^{\nu'} q_\rho^{\rho'} \nabla_{\mu'} D_{\nu'} u_{\rho'} \\
&= q_\mu^{\mu'} q_\nu^{\nu'} q_\rho^{\rho'} \nabla_{\mu'} q_{\nu'}^{\nu''} q_{\rho'}^{\rho''} \nabla_{\nu''} u_{\rho''} \qquad \text{(I.1.1. 7)}
\end{aligned}
$$

The outer derivative hits either a q tensor or ∇u the latter of which will give rise to a curvature term. Consider then the ∇q terms.
Since ∇ is g compatible we have $\nabla q = s\nabla n \otimes n = s[(\nabla n) \otimes n + n \otimes (\nabla n)]$. Since all of these terms are contracted with q tensors and q annihilates n, the only terms that survive are proportional to terms of either the form

$$
(\nabla_{\mu'} n_{\nu'})(n^{\rho''}(\nabla_{\nu''} u_{\rho''}) = -(\nabla_{\mu'} n_{\nu'})(\nabla_{\nu''} n^{\rho''}) u_{\rho''}
$$

where $n^\mu u_\mu = 0 \Rightarrow \nabla_\nu (n^\mu u_\mu) = 0$ was exploited, or of the form $(\nabla_{\mu'} n_{\nu'})(\nabla_n u_{\rho'})$. Concluding, the only terms that survive from ∇q terms can be transformed terms into proportional to $\nabla n \otimes \nabla n$ or $\nabla n \otimes \nabla_n u$ where the ∇n factors, since contracted with q tensors, can be traded for extrinsic curvature terms (use $u_\mu = q_\mu^\nu u_\nu$ to do that).
It turns out that the terms proportional to $\nabla_n u$ cancel each other when computing the antisymmetrized second D derivative of u due to the symmetry of K and we are thus left with the famous

Gauss equation

$$\begin{aligned}
R^{(D)\ \sigma}_{\mu\nu\rho} u_\sigma : &= 2D_{[\mu}D_{\nu]}u_\rho \\
&= [2sK_{\rho[\mu}K^\sigma_{\nu]} + q^{\mu'}_\mu q^{\nu'}_\nu q^{\rho'}_\rho q^\sigma_{\sigma'} R^{(D+1)\ \sigma'}_{\mu'\nu'\rho'}]u_\sigma \\
R^{(D)}_{\mu\nu\rho\sigma} &= 2sK_{\rho[\mu}K_{\nu]\sigma} + q^{\mu'}_\mu q^{\nu'}_\nu q^{\rho'}_\rho q^{\sigma'}_\sigma R^{(D+1)}_{\mu'\nu'\rho'\sigma'} \qquad \text{(I.1.1. 8)}
\end{aligned}$$

Using this general formula we can specialize to the Riemann curvature scalar which is our ultimate concern in view of the Einstein-Hilbert action. Employing the standard abbreviations $K := K_{\mu\nu}q^{\mu\nu}$ and $K^{\mu\nu} = q^{mu\rho}q^{\nu\sigma}K_{\nu\sigma}$ (notice that indices for spatial tensors can be moved either with q or with g) we obtain

$$\begin{aligned}
R^{(D)} &= R^{(D)}_{\mu\nu\rho\sigma}q^{\mu\rho}q^{\nu\sigma} \\
&= s[K^2 - K_{\mu\nu}K^{\mu\nu}] + q^{\mu\rho}q^{\nu\sigma}R^{(D+1)}_{\mu\nu\rho\sigma} \qquad \text{(I.1.1. 9)}
\end{aligned}$$

Equation (I.1.1. 9) is not yet quite what we want since it is not yet purely expressed in terms of $R^{(D+1)}$ alone. However, we can eliminate the second term in (I.1.1. 9) by using $g = q + sn \otimes n$ and the definition of curvature $R^{(D+1)}_{\mu\nu\rho\sigma}n^\sigma = 2\nabla[\mu\nabla_{\nu]}n_\rho$ as follows

$$\begin{aligned}
R^{(D+1)} &= R^{(D+1)}_{\mu\nu\rho\sigma}g^{\mu\rho}g^{\nu\sigma} \\
&= q^{\mu\rho}q^{\nu\sigma}R^{(D+1)}_{\mu\nu\rho\sigma} + 2sq^{\rho\mu}n^\nu[\nabla_\mu, \nabla_\nu]n_\rho \\
&= q^{\mu\rho}q^{\nu\sigma}R^{(D+1)}_{\mu\nu\rho\sigma} + 2sn^\nu[\nabla_\mu, \nabla_\nu]n_\nu \qquad \text{(I.1.1. 10)}
\end{aligned}$$

where in the first step we used the antisymmetry of the Riemann tensor to eliminate the term quartic in n and in the second step we used again $q = g - sn \otimes n$ and the antisymmetry in the $\mu\nu$ indices. Now

$$n^\nu([\nabla_\mu, \nabla_\nu]n^\mu) = -(\nabla_\mu n^\nu)(\nabla_\nu n^\mu) = +(\nabla_\mu n^\mu)(\nabla_\nu n^\nu) + \nabla_\mu(n^\nu\nabla_\nu n^\mu - n^\mu\nabla_\nu n^\nu)$$

and using $\nabla_\mu s = 2n^\nu\nabla_\mu n_\nu = 0$ we have

$$\begin{aligned}
\nabla_\mu n^\mu &= g^{\mu\nu}\nabla_\nu n^\mu = q^{\mu\nu}\nabla_\nu n^\mu = K \qquad \text{(I.1.1. 11)} \\
&(\nabla_\mu n^\nu)(\nabla_\nu n^\mu) = g^{\nu\sigma}g^{\rho\mu}(\nabla_\mu n_\sigma)(\nabla_\nu n_\rho) = q^{\nu\sigma}q^{\rho\mu}(\nabla_\mu n_\sigma)(\nabla_\nu n_\rho) = K_{\mu\nu}K^{\mu\nu}
\end{aligned}$$

Combining (I.1.1. 9), (I.1.1. 10) and (I.1.1. 11) we obtain the *Codacci equation*

$$R^{(D+1)} = R^{(D)} - s[K_{\mu\nu}K^{\mu\nu} - K^2] + 2s\nabla_\mu(n^\nu\nabla_\nu n^\mu - n^\mu\nabla_\nu n^\nu) \qquad \text{(I.1.1. 12)}$$

Inserting this differential geometric identity back into the action, the third term in (I.1.1. 12) is a total differential which we drop for the time being as one can rederive it later on when making the variational principle well-defined.

At this point it is useful to pull back various quantities to σ. Consider the D spatial vextor fields on Σ_t defined by

$$X^\mu_a(X) := X^\mu_{,a}(x,t)_{|X(x,t)=X} \qquad \text{(I.1.1. 13)}$$

Then we have due to $n_\mu X^\mu_a = 0$ that

$$q_{ab}(t,x) := (X^\mu_{,a}X^\nu_{,b}q_{\mu\nu})(X(x,t)) = g_{\mu\nu}(X(t,x))X^\mu_{,a}(t,x)X^\nu_{,b}(t,x) \qquad \text{(I.1.1. 14)}$$

and

$$K_{ab}(t,x) := (X^\mu_{,a}X^\nu_{,b}K_{\mu\nu})(X(x,t)) = (X^\mu_{,a}X^\nu_{,b}\nabla_\mu n_\nu)(t,x) \qquad \text{(I.1.1. 15)}$$

Using q_{ab} and its inverse $q^{ab}=\epsilon^{aa_1..a_{D-1}}\epsilon^{bb_1..b_{D-1}}q_{a_1b_1}..q_{a_{D-1}b_{D-1}}/[\det((q_{cd}))\ (D-1)!]$ we can express $q_{\mu\nu}, q^{\mu\nu}, q_\mu^\nu$ as

$$
\begin{aligned}
q^{\mu\nu}(X) &= [q^{ab}(x,t)X^\mu_{,a}X^\mu_{,b}](x,t)_{|X(x,t)=X} \\
q^\nu_\mu(X) &= g_{\mu\rho}(X)q^{\rho\nu}(X) \\
q_{\mu\nu}(X) &= g_{\nu\rho}(X)q^\rho_\mu(X)
\end{aligned}
\tag{I.1.1. 16}
$$

To verify that this coincides with our previous defintion $q=g-sn\otimes n$ it is sufficient to check the matrix elements in the basis given by the vector fields n, X_a. Since for both definitions n is annihilated we just need to verify that (I.1.1. 16) when contracted with $X_a\otimes X_b$ reproduces (I.1.1. 14) which is indeed the case.

Next we define $N(x,t):=N(X(x,t)), \vec{N}^a(x,t):=q^{ab}(x,t)(X^\mu_b g_{\mu\nu}N^\nu)(X(x,t))$. Then it is easy to verify that

$$
K_{ab}(x,t)=\frac{1}{2N}(\dot{q}_{ab}-(\mathcal{L}_{\vec{N}}q)_{ab})(x,t) \tag{I.1.1. 17}
$$

We can now pull back the expressions quadratic in $K_{\mu\nu}$ that appear in (I.1.1. 12) using (I.1.1. 16) and find

$$
\begin{aligned}
K(x,t) &= (q^{\mu\nu}K_{\mu\nu})(X(x,t))=(q^{ab}K_{ab})(x,t) \\
(K_{\mu\nu}K^{\mu\nu})(x,t) &= (K_{\mu\nu}K_{\rho\sigma}q^{\mu\rho}q^{\nu\sigma})(X(x,t))=(K_{ab}K_{cd}q^{ac}q^{bd})(x,t)
\end{aligned}
\tag{I.1.1. 18}
$$

Likewise we can pull back the curvature scalar $R^{(D)}$. We have

$$
R^{(D)}(x,t)=(R^{(D)}_{\mu\nu\rho\sigma}q^{\mu\rho}q^{\nu\sigma})(X(x,t))(R^{(D)}_{\mu\nu\rho\sigma}X^\mu_aX^\nu_bX^\rho_cX^\sigma_d)(X(x,t))q^{ac}(x,t)q^{bd}(x,t) \tag{I.1.1. 19}
$$

We would like to show that this expression equals the curvature scalar R as defined in terms of the Christoffel connection for q_{ab}. To see this it is sufficient to compute $(X^\mu_a D_\mu f)(X(x,t))=\partial_a f(X(x,t))=:(D_af)(x,t)$ with $f(x,t):=F(X(x,t))$ and with $u_a(x,t):=(X^\mu_a u_\mu)(X(x,t)),\ u^a(x,t)=q^{ab}(x,t)u_b(x,t)$

$$
\begin{aligned}
(D_au_b)(x,t) &:= (X^\mu_aX^\nu_bD_\mu u_\nu)(X(x,t)) \\
&= X^\mu_{,a}(x,t)X^\nu_{,b}(x,t)(\nabla_\mu u_n u)(X(x,t)) \\
&= (\partial_a u_b)(x,t)-X^\mu_{,ab}u_\mu(X(x,t)) \\
&\quad -u^c(x,t)\Gamma^{(D+1)}_{\rho\mu\nu}(X(x,t))X^\rho_{,c}(x,t)(X^\mu_{,a}(x,t)X^\nu_{,b}(x,t) \\
&= (\partial_a u_b)(x,t)-\Gamma^{(D)}_{cab}(x,t)u^c(x,t)
\end{aligned}
\tag{I.1.1. 20}
$$

where in the last step we have used the explicit expressions of the Christoffel connections $\Gamma^{(D+1)}$ and $\Gamma^{(D)}$ in terms of $g_{\mu\nu}$ and q_{ab} respectively. Now since every tensor field W is a linear combination of tensor products of one forms and since D_μ satisfies the Leibniz rule we easily find $(X^\mu_aX^\nu_b..D_\mu W_{\nu..})(X(x,t))=:(D_aW_{b..})(x,t)$ where now D_a denotes the uniqe torsion-free covariant differential associated with q_{ab} and $W_{a..}$ is the pull-back of $W_{\mu..}$. In particular, we have $X^\mu_aX^\nu_bX^\rho_cD_\mu D_\nu u_\rho=D_aX^\mu_bX^\nu_cD_\mu u_\rho=D_aD_bu_c$ from which our assertion follows since

$$
\begin{aligned}
(R_{abcd}u_d)(x,t):=([D_a,D_b]u_c)(x,t) &= (X^\mu_aX^\nu_bX^\rho_c[D_\mu,D_\nu]u_\rho)(X(x,t)) \\
&= (X^\mu_aX^\nu_bX^\rho_cX^\sigma_dR^{(D)}_{\mu\nu\rho\sigma})(X(x,t))u^d(x,t)
\end{aligned}
\tag{I.1.1. 21}
$$

From now on we will move indices with the metric q_{ab} only.

One now expresses the line element in the new system of coordinates x, t using the quantities q_{ab}, N, N^a (we refrain from displaying the arguments of the components of the metric)

$$\begin{aligned} ds^2 &= g_{\mu\nu} dX^\mu \otimes dX^\nu \\ &= g_{\mu\nu}(X(t,x))[X^\mu_{,t} dt + X^\mu_{,a} dx^a] \otimes [X^\nu_{,t} dt + X^\nu_{,b} dx^b] \\ &= g_{\mu\nu}(X(t,x))[N n^\mu dt + X^\mu_{,a}(dx^a + N^a dt)] \otimes [N n^\nu dt + X^\nu_{,b}(dx^b + N^b dt)] \\ &= [sN^2 + q_{ab}N^a N^b] dt \otimes dt + q_{ab}N^b[dt \otimes dx^a + dx^a \otimes dt] + q_{ab} dx^a \otimes dx^b \end{aligned} \tag{I.1.1. 22}$$

and reads off the components g_{tt}, g_{ta}, g_{ab} of $X^* g$ in this frame. Since the volume form $\Omega(X) := \sqrt{|\det(g)|} d^{D+1}X$ is covariant, i.e., $(X^*\Omega)(x,t) = \sqrt{|\det(X * g)|} dt d^D x$ we just need to compute $\det(X^* g) = sN^2 \det(q_{ab})$ in order to finally cast the action (I.1.1. 1) into $D+1$ form. The result is (dropping the total differential in (I.1.1. 12))

$$S = \frac{1}{\kappa} \int_{\mathbb{R}} dt \int_\sigma d^D x \sqrt{\det(q)} |N| (R - s[K_{ab}K^{ab} - (K_a^a)^2]) \tag{I.1.1. 23}$$

We could drop the absolute sign for N in (I.1.1. 23) since we took N positive but we will keep it for the moment to see what happens if we allow arbitrary sign. Notice that (I.1.1. 23) *vanishes identically* for $D = 1$, indeed in two spacetime dimensions the Einstein action is proportional to a topological charge, the so-called Euler characteristic of M and in what follows we concentrate on $D > 1$.

We now wish to cast this action into canonical form, that is, we would like to perform the Legendre transform from the Lagrangean density appearing in (I.1.1. 23) to the corresponding Hamiltonian density. The action (I.1.1. 23) depends on the velocities $\dot{q}_{ab}$ of q_{ab} but not on those of N and N^a. Therefore we obtain for the conjugate momenta (use (I.1.1. 17) and the fact that R does not contain time derivatives)

$$\begin{aligned} \frac{1}{\kappa} P^{ab}(t,x) &:= \frac{\delta S}{\delta \dot{q}_{ab}(t,x)} = -s\frac{|N|}{N\kappa}\sqrt{\det(q)}[K^{ab} - q^{ab}(K_c^c)] \\ \Pi(t,x) &:= \frac{\delta S}{\delta \dot{N}(t,x)} = 0 \\ \Pi_a(t,x) &:= \frac{\delta S}{\delta \dot{N}^a(t,x)} = 0 \end{aligned} \tag{I.1.1. 24}$$

The Lagrangean in (I.1.1. 23) is therefore a *singular* Lagrangean, one cannot solve all velocities for momenta [51]. We can solve $\dot{q}_{ab}$ in terms of q_{ab}, N, N^a and P^{ab} using (I.1.1. 17) but this is not possible for $\dot{N}, \dot{N}^a$, rather we have the so-called *primary constraints*

$$C(t,x) := \Pi(t,x) = 0 \text{ and } C^a(t,x) := \Pi^a(t,x) = 0 \tag{I.1.1. 25}$$

The Hamiltonian treatment of systems with constraints has been developed by Dirac [24] to which a short introduction is given in section III.1. According to that theory, we are supposed to introduce Lagrange multiplier fields $\lambda(t,x), \lambda_a(t,x)$ for the primary constraints and to perform the Legendre transform as usual with respect to the remaining velocities which can be solved for. We have

$$\begin{aligned} \dot{q}_{ab} &= 2NK_{ab} + (\mathcal{L}_{\vec{N}} q)_{ab} \\ \dot{q}_{ab}P^{ab} &= (\mathcal{L}_{\vec{N}} q)_{ab}P^{ab} - 2s|N|\sqrt{\det(q)}[K_{ab}K^{ab} - K^2] \\ P_{ab}P^{ab} &= \det(q)(K_{ab}K^{ab} + (D-2)K^2) \\ P^2 &:= (P_a^a)^2 = (1-D)^2 \det(q) K^2 \end{aligned} \tag{I.1.1. 26}$$

and by means of these formulae we obtain the canonical form of the action (I.1.1. 23)

$$\begin{aligned}
\kappa S &= \int_{\mathbb{R}} dt \int_\sigma d^D x\{\dot{q}_{ab}P^{ab} + \dot{N}\Pi + \dot{N}^a\Pi_a \qquad\qquad (\text{I.1.1. } 27)\\
&-[\dot{q}_{ab}(P,q,N,\vec{N})P^{ab} + \lambda C + \lambda^a C_a - \sqrt{\det(q)}|N|(R - s[K_{ab}K^{ab} - K^2])(P,q,N,\vec{N})]\}\\
&= \int_{\mathbb{R}} dt \int_\sigma d^D x\{\dot{q}_{ab}P^{ab} + \dot{N}\Pi + \dot{N}^a\Pi_a\\
&-[(\mathcal{L}_{\vec{N}}q)_{ab}P^{ab} + \lambda C + \lambda^a C_a - \sqrt{\det(q)}|N|(R + s[K_{ab}K^{ab} - K^2])(P,q,N,\vec{N})]\}\\
&= \int_{\mathbb{R}} dt \int_\sigma d^D x\{\dot{q}_{ab}P^{ab} + \dot{N}\Pi + \dot{N}^a\Pi_a\\
&-[(\mathcal{L}_{\vec{N}}q)_{ab}P^{ab} + \lambda C + \lambda^a C_a + |N|(-\frac{s}{\sqrt{\det(q)}}[P_{ab}P^{ab} - \frac{1}{D-1}P^2] - \sqrt{\det(q)}R)]\}
\end{aligned}$$

Upon performing a spatial integration by parts (whose boundary term vanishes since $\partial\sigma = \emptyset$) one can cast it into the following more compact form

$$S = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\sigma d^D x\{\dot{q}_{ab}P^{ab} + \dot{N}\Pi + \dot{N}^a\Pi_a - [\lambda C + \lambda^a C_a + N^a H_a + |N|H]\} \qquad (\text{I.1.1. } 28)$$

where

$$\begin{aligned}
H_a &:= -2q_{ac}D_b P^{bc}\\
H &:= -(\frac{s}{\sqrt{\det(q)}}[q_{ac}q_{bd} - \frac{1}{D-1}q_{ab}q_{cd}]P^{ab}P^{cd} + \sqrt{\det(q)}R) \qquad (\text{I.1.1. } 29)
\end{aligned}$$

are called the *(spatial) Diffeomorphism constraint* and *Hamiltonian constraint* respectively, for reasons which we will derive below.

The geometrical meaning of these quantities is as follows :
At fixed t the fields $(q_{ab}(t,x), N^a(t,x), N(t,x); P^{ab}(x,t), \Pi_a(t,x), \Pi(t,x))$ label points (configuration;canonically conjugate momenta) in an infinite dimensional phase space $\mathcal{M}$ (or symplectic manifold). Strictly speaking, we should now specify on what Banach space this manifold is modelled [52], however, we will be brief here as we are primarily not interested in the metric formulation of this section but rather in the connection formulation of the next section where we will give more details. For the purpose of this subsection it is sufficient to say that we can choose the model space to be the direct product of the space $T_2(\sigma)\times T_1(\sigma)\times T_0(\sigma)$ of smooth symmetric covariant tensor fileds of rank $2,1,0$ on σ respectively and the space $\tilde{T}^2(\sigma)\times\tilde{T}^1(\sigma)\times\tilde{T}^0(\sigma)$ of smooth symmetric contravariant tensor density fields of weight one and of rank $2,1,0$ on σ respectively, equipped with some Sobolev norm. (The precise functional analytic description is somewhat more complicated in case that σ is unbounded with boundary but can also be treated). In particular, one shows that the action (I.1.1. 28) is differentiable in this topology.

The phase space carries the strong [52] symplectic structure Ω or Poisson bracket

$$\{P(f^2), F_2(q)\} = \kappa F_2(f^2),\ \{\vec{\Pi}(\vec{f}^1), \vec{F}_1(\vec{N})\} = \kappa\vec{F}_1(\vec{f}^1),\ \{\Pi(f), F(N)\} = \kappa F(f) \qquad (\text{I.1.1. } 30)$$

(all other brackets vanishing) where we have defined the following pairing, invariant under diffeomorphisms of σ, e.g.

$$\tilde{T}^2(\sigma)\times T_2(\sigma)\to\mathbb{R};\ (F_2, f^2)\to F^2(f_2) := \int_\sigma d^D x F_2^{ab}(x) f^2_{ab}(x) \qquad (\text{I.1.1. } 31)$$

and similar for the other fields. Physicists use the following short-hand notation for (I.1.1. 30)

$$\{P^{ab}(t,x), q_{cd}(t,x')\} = \kappa\delta^a_{(c}\delta^b_{d)}\delta^{(D)}(x,y) \ . \tag{I.1.1. 32}$$

In the language of symplectic geometry, the first term in the action (I.1.1. 28) is a symplectic potential for the symplectic structure (I.1.1. 30). We now turn to the meaning of the term in the square bracket in (I.1.1. 28), that is, the "Hamiltonian"

$$\kappa\mathbf{H} := \int_\sigma d^Dx[\lambda C + \lambda^a C_a + N^a H_a + |N|H] =: \vec{C}(\vec{\lambda}) + C(\lambda) + \vec{H}(\vec{N}) + H(|N|) \tag{I.1.1. 33}$$

of the action and the associated equations of motion.

The variation of the action with respect to the Lagrange multiplier fields $\vec{\lambda}, \lambda$ reproduces the primary constraints (I.1.1. 25). If the dynamics of the system is to be consistent, then these constraints must be preserved under the evolution of the system, that is, we should have e.g. $\dot{C}(t,x) := \{\mathbf{H}, C(t,x)\} = 0$ for all $x \in \sigma$, or equivalently, $\dot{C}(f) := \{\mathbf{H}, C(f)\} = 0$ for all (t-independent) smearing fields $f \in T_0(\sigma)$. However, we do not get zero but rather

$$\{\vec{C}(\vec{f}), \mathbf{H}\} = \vec{H}(\vec{f}) \text{ and } \{C(f), \mathbf{H}\} = H((\frac{N}{|N|}f) \tag{I.1.1. 34}$$

which is supposed to vanish for all $f, \vec{f}$. Thus, consistency of the equations of motion ask us to impose the *secondary constraints*

$$H(x,t) = 0 \text{ and } H_a(x,t) = 0 \tag{I.1.1. 35}$$

for all $x \in \sigma$. Since these two functions appear next to the C, C_a in (I.1.1. 33), in general relativity the "Hamiltonian" is constrained to vanish ! General relativity is an example of a so-called constrained Hamiltonian system with no true Hamiltonian. The reason for this will become evident in a moment.

Now one might worry that imposing consistency of the secondary constraints under evolution results in tertiary constraints etc., but fortunately, this is not the case. Consider the smeared quantities $H(f), \vec{H}(\vec{f})$ where, e.g., $\vec{H}(\vec{N}) := \int_\sigma d^3x N^a V_a$ (notice that indeed H, Π and H_a, Π_a are, respectively, scalar and co-vector densities of weight one on σ). Then we obtain

$$\begin{aligned} \{\mathbf{H}, \vec{H}(\vec{f})\} &= \vec{H}(\mathcal{L}_{\vec{N}}\vec{f}) - H(\mathcal{L}_{\vec{f}}|N|) \\ \{\mathbf{H}, H(f)\} &= H(\mathcal{L}_{\vec{N}}f) + \vec{H}(\vec{N}(|N|, f, q)) \end{aligned} \tag{I.1.1. 36}$$

where $\vec{N}(f,f',q)^a = q^{ab}(ff'_{,b} - f'f_{,b})$. Equations (I.1.1. 36) are equivalent to the *Dirac algebra* [24]

$$\begin{aligned} \{\vec{H}(\vec{f}), \vec{H}(\vec{f'})\} &= \kappa\vec{H}(\mathcal{L}_{\vec{f}}\vec{f'}) \\ \{\vec{H}(\vec{f}), H(f))\} &= \kappa H(\mathcal{L}_{\vec{f}}f) \\ \{H(f), H(f'))\} &= \kappa\vec{H}(\vec{N}(f,f',q)) \end{aligned} \tag{I.1.1. 37}$$

also called the *hypersurface deformation algebra*. The meaning of (I.1.1. 34,I.1.1. 37) is that the constraint surface $\overline{\mathcal{M}}$ of $\mathcal{M}$, the submanifold of $\mathcal{M}$ where the constraints hold, is preserved under the motions generated by the constraints. In the terminology of Dirac [24], all constraints are of first class (determine coisotropic constraint submanifolds [51] of $\mathcal{M}$) rather than of second class (determine symplectic constraint submanifolds [51] of $\mathcal{M}$).

It remains to study the equations of motion of the canonical coordinates on the phase space themselves. Since $C = \Pi, C_a = \Pi_a$ it remains to study those of N, N^a, q_{ab}, P^{ab}. For shift and lapse we obtain $\dot{N}^a = \lambda^a, \dot{N} = \lambda$. Since λ^a, λ are *arbitrary, unspecified functions* we see that also the trajectory of lapse and shift is completely arbitrary. Moreover, the equations of motion of q_{ab}, P^{ab} are completely unaffected by the term $\vec{C}(\vec{\lambda})+C(\lambda)$ in $\mathbf{H}$. It is therefore completely straightforward to solve the equations of motion as far as N, N^a, Π, Π_a are concerned : Simply treat N, N^a as *Lagrange multipliers* and drop all terms proportional to C, C_a from the action (I.1.1. 28). The result is the reduced action

$$S = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\sigma d^Dx\{\dot{q}_{ab}P^{ab} - [N^a H_a + |N|H]\} \tag{I.1.1. 38}$$

called the *Arnowitt – Deser – Misner action* [46]. It is straightforward to check that as far as q_{ab}, P^{ab} are concerned, the actions (I.1.1. 28) and (I.1.1. 38) are completely equivalent.

The equations of motion of q_{ab}, P^{ab} then finally allow us to interpret the motions that the constraints generate on $\mathcal{M}$ geometrically. Since the reduced Hamiltonian (using the same symbol as before)

$$\mathbf{H} = \frac{1}{\kappa}\int_\sigma d^Dx[N^a H_a + |N|H] \tag{I.1.1. 39}$$

is a linear combination of constraints we obtain the equations of motion once we know the Hamiltonian flow of the functions $\vec{H}(\vec{f}), H(f)$ for any $\vec{f}, f$ separately. Denoting, for any function J on $\mathcal{M}$,

$$\delta_{\vec{f}}J := \{\vec{H}(\vec{f}), J\} \text{ and } \delta_f J := \{H(f), J\} \tag{I.1.1. 40}$$

it is easiest to begin with the corresponding equations for $J = F_2(q)$ since upon integration by parts we have $\vec{H}(\vec{f}) = \int d^DxP^{ab}(\mathcal{L}_{\vec{f}}q)_{ab}$ so that both constraint functions are simple polynomials in P^{ab} not involving their derivatives. We then readily find

$$\begin{aligned}
\delta_{\vec{f}}F_2(q) &= \kappa F_2(\mathcal{L}_{\vec{f}}q)\\
\delta_f F_2(q) &= -2s\kappa\int_\sigma d^Dx\frac{P_{ab} - Pq_{ab}/(D-1)}{\sqrt{\det(q)}}
\end{aligned} \tag{I.1.1. 41}$$

Using the relations (I.1.1. 24), (I.1.1. 17) the second identity in (I.1.1. 41) can be written as

$$\delta_{|N|}q_{ab} = 2N\kappa K_{ab} = \kappa(\dot{q}_{ab} - (\mathcal{L}_{\vec{N}}q)_{ab})$$

In order to interprete this quantity, notice that the components of n_μ in the frame t, x^a are given by $n_t = n_\mu X^\mu_{,t} = sN,\ n_a = n_\mu X^\mu_{,a} = 0$. In order to compute the contravariant components n^μ in that frame we need the corresponding contravariant metric components. From (I.1.1. 22) we find the covariant components to be $g_{tt} = sN^2 + q_{ab}N^aN^b, g_{ta} = q_{ab}N^b, g_{ab} = q_{ab}$ so that the inverse metric has components $g^{tt} = s/N^2, g^{ta} = -sN^a/N^2, g^{ab} = q^{ab} + sN^aN^b/N^2$. Thus $n^t = 1/N, n^a = -N^a/N$ and since $q_{at} = q_{tt} = 0$ we finally obtain

$$\delta_{|N|}F_2(q) = \kappa F_2(\mathcal{L}_{Nn}q) \tag{I.1.1. 42}$$

which of course we guessed immediately from the $D+1$ dimensional identiy (I.1.1. 6). Concluding, as far as q_{ab} is concerned, H_a generates *on all of* $\mathcal{M}$ diffeomorphisms of M that preserve Σ_t while H generates diffeomorphisms of M orthogonal to Σ_t.

The corresponding computation for $P(f^2)$ is harder by an order of magnitude due to the curvature term involved in H and due to the fact that the identity corresponding to (I.1.1. 42) holds *only on* *shell*, that is, when the (vacuum) Einstein equations $G^{(D+1)}_{\mu\nu} := R^{(D+1)}_{\mu\nu} - \frac{g_{\mu\nu}}{2} R^{(D+1)} = 0$ hold. The variation with respect to $\vec{H}(\vec{f}) = -\int_\sigma d^D x q_{ab} (\mathcal{L}_{\vec{f}} P)^{ab}$ (notice that P^{ab} carries density weight one to verify this identity) is still easy and yields the expected result

$$\delta_{\vec{f}} P(f^2) = \kappa(\mathcal{L}_{\vec{f}} P)(f^2) \qquad \text{(I.1.1. 43)}$$

We will now describe the essential steps for the analog of (I.1.1. 42). The ambitious reader who wants to fill in the missing steps should expect to perform at least one Din A4 page of calculation in between each of the subsequent formulae.
We start from formula (I.1.1. 29). Then

$$\begin{aligned}
\{H(|N|), P^{ab}\} &= \frac{\delta H(|N|)}{\delta q_{ab}} \\
&= \frac{s|N|}{\sqrt{\det(q)}}[2(P^{ac}P_c^b - P^{ab}P/(D-1)) - \frac{q^{ab}}{2}(P^{cd}P_{cd} - P^2/(D-1))] \\
&\quad + \frac{\delta}{\delta q_{ab}} \int d^D x |N| \sqrt{\det(q)} R
\end{aligned} \qquad \text{(I.1.1. 44)}$$

where the second term comes from the $\sqrt{\det(q)}^{-1}$ factor and we used the well-known formula $\delta \det(q) = \det(q) q^{ab} \delta q_{ab}$. To perform the remaining variation in (I.1.1. 44) we write

$$\delta\sqrt{\det(q)}R = [\delta\sqrt{\det(q)}]R + \sqrt{\det(q)}[\delta q^{ab}]R_{ab} + \sqrt{\det(q)}q^{ab}[\delta R_{ab}]$$

use $\delta\delta^a_b = \delta[q^{ac}q_{cb}] = 0$ in the second variation and can simplify (I.1.1. 44)

$$\begin{aligned}
\{H(|N|), P^{ab}\} &= \frac{2s|N|}{\sqrt{\det(q)}}[2(P^{ac}P_c^b - P^{ab}P/(D-1)] + \frac{q^{ab}|N|H}{2} + |N|\sqrt{\det(q)}(q^{ab}R - R^{ab}) \\
&\quad + \int d^D x |N| \sqrt{\det(q)} q^{cd} \frac{\delta}{\delta q_{ab}} R_{cd}
\end{aligned} \qquad \text{(I.1.1. 45)}$$

The final variation is the most difficult one since R_{cd} contains second derivatives of q_{ab}. Using the explicit expression of R_{abcd} in terms of the Christoffel connection Γ^c_{ab} and observing that, while the connection itself is not a tensor, its variation in fact *is* a tensor, we find after careful use of the definition of the covariant derivative

$$q^{cd}\delta R_{cd} = q^{cd}[-D_c\delta\Gamma^e_{ed} + D_e\delta\Gamma^e_{cd}] \qquad \text{(I.1.1. 46)}$$

We now use the explicit expression of Γ^a_{bc} in terms of q_{ab} and find

$$\delta\Gamma^a_{bc} = \frac{q^{ad}}{2}[D_c\delta q_{bd} + D_b\delta q_{cd} - D_d\delta q_{bc}] \qquad \text{(I.1.1. 47)}$$

Next we insert (I.1.1. 46) and (I.1.1. 47) into the integral appearing in (I.1.1. 45) and integrate by parts two times using the fact that for the divergence of a vector v^a we have $\sqrt{\det(q)}D_a v^a = D_a(\sqrt{\det(q)}v^a) = \partial_a(\sqrt{\det(q)}v^a)$ (no boundary terms due to $\partial\sigma = \emptyset$) and find

$$\begin{aligned}
\int d^D x |N| \sqrt{\det(q)} q^{cd}\delta R_{cd} &= \int d^D x \sqrt{\det(q)} q^{cd}[(D_c|N|)\delta\Gamma^e_{ed} - (D_e|N|)\delta\Gamma^e_{cd}] \\
&= \int d^D x \sqrt{\det(q)} q^{cd} q^{ef}[(D_c|N|)(D_d\delta q_{ef}) - (D_e|N|)(D_c\delta q_{df})] \\
&= \int d^D x \sqrt{\det(q)}[-(D_c D^c|N|)q^{ab} + (D^a D^b|N|)]\delta q_{ab}
\end{aligned} \qquad \text{(I.1.1. 48)}$$

Collecting all contributions we obtain the desired result

$$\{H(|N|),P^{ab}\}=\frac{2s|N|}{\sqrt{\det(q)}}[2(P^{ac}P_c^b-P^{ab}P/(D-1)]+\frac{q^{ab}|N|H}{2}$$
$$+|N|\sqrt{\det(q)}(q^{ab}R-R^{ab})+\sqrt{\det(q)}[-(D_cD^c|N|)q^{ab}-(D^aD^b|N|)] \qquad \text{(I.1.1. 49)}$$

which does not look at all as $\mathcal{L}_{Nn}P^{ab}$!

In order to compute $\mathcal{L}_{Nn}P^{ab}$ we need an identity for $\mathcal{L}_{Nn}K_{\mu\nu}=N\mathcal{L}_nK_{\mu\nu}$ which we now derive. Using the definition of the Lie derivative in terms of the covariant derivative ∇_μ and using $g=q+sn\otimes n$ one finds first of all

$$\mathcal{L}_nK_{\mu\nu}=-KK_{\mu\nu}+2K_{\rho\mu}K_\nu^\rho+[\nabla_\rho(n^\rho K_{\mu\nu})+2sK_{\rho(\mu}n_{\nu)}\nabla_nn^\rho] \qquad \text{(I.1.1. 50)}$$

Using the Gauss equation (I.1.1. 8) we find for the Ricci tensor $R_{\mu\nu}^{(D)}$ the following equation (use again $g=q+sn\otimes n$ and the defintion of curvature as $R=[\nabla,\nabla]$)

$$R_{\rho\sigma}^{(D+1)}q_\mu^\rho q_\nu^\sigma-R_{\mu\nu}^{(D)}=s[-K_{\mu\nu}K+K_{\mu\rho}K_\nu^\rho+q_\mu^\rho q_\nu^\sigma n^\lambda[\nabla_\rho,\nabla_\lambda]n_\sigma] \qquad \text{(I.1.1. 51)}$$

We claim that the term in square brackets on the right hand side of (I.1.1. 50) equals $(-s)$ times the sum of the left hand side of (I.1.1. 51) and the term $-s(D_\mu D_\nu N)/N$. In order to prove this we manipulate the commutator of covariant derivatives appearing in (I.1.1. 51) making use of the definition of the extrinsic curvature. One finds

$$\begin{aligned}&q_\mu^\rho q_\nu^\sigma n^\lambda[\nabla_\rho,\nabla_\lambda]n_\sigma]\\ =\;&q_\mu^\rho q_\nu^\sigma n^\lambda(\nabla_\rho\nabla_\lambda n_\sigma)+KK_{\mu\nu}-\nabla_\rho(n^\rho K_{\mu\nu})\\ &-s(\nabla_nn^\rho)n_\nu K_{\mu\rho}-s(\nabla_n(n_\mu n^\rho))(\nabla_\rho n_\nu)\end{aligned} \qquad \text{(I.1.1. 52)}$$

Using this identity we find for the sum of the term in square brackets on the right hand side of (I.1.1. 50) and s times the sum of the right hand side of (I.1.1. 51) the expression (dropping the obvious cancellations)

$$\begin{aligned}&K_{\mu\rho}K_\nu^\rho+q_\mu^\rho q_\nu^\sigma n^\lambda(\nabla_\rho\nabla_\lambda n_\sigma)+s[K_{\rho\nu}n_\mu(\nabla_nn^\rho)-(\nabla_n(n_\mu n^\rho))(\nabla_\rho n_\nu)]\\ =\;&K_{\mu\rho}K_\nu^\rho+q_\mu^\rho q_\nu^\sigma n^\lambda(\nabla_\rho\nabla_\lambda n_\sigma)+s[n_\mu(\nabla_nn^\rho)\{q_\rho^\sigma-\delta_\rho^\sigma\}(\nabla_\sigma n_\nu)-(\nabla_nn\mu)(\nabla_nn_\nu)]\\ =\;&K_{\mu\rho}K_\nu^\rho+q_\mu^\rho q_\nu^\sigma(\nabla_\rho\nabla_nn_\sigma)-q_\mu^\rho q_\nu^\sigma(\nabla_\rho n^\lambda)(\nabla_\lambda n_\sigma)-s(\nabla_nn\mu)(\nabla_nn_\nu)\\ =\;&+q_\mu^\rho q_\nu^\sigma(\nabla_\rho\nabla_nn_\sigma)-s(\nabla_nn\mu)(\nabla_nn_\nu)\end{aligned} \qquad \text{(I.1.1. 53)}$$

where in the second step it has been used that the curly bracket vanishes since it is proportional to n_ρ and contracted with the spatial vector ∇_nn^ρ, in the third step we moved ∇^λ inside a covariant derivative and picked up a correction term and in the fourth step one realizes that this correction term is just the negative of the first term using that $K_{\mu\nu}=q_\mu^\rho\nabla_\rho n_\nu$. Our claim is equivalent to showing that the last line of (I.1.1. 53) is indeed given by $-s(D_\mu D_\nu N)/N$.
To see this notice that if the surface Σ_t is defined by $t(X)=t=const.$ then $1=T^\mu\nabla_\mu t$. Since $\nabla_\mu t$ is orthogonal to Σ_t we have $n_\mu=sN\nabla_\mu t$ as one verifies by contracting with T^μ and thus $N=1/(\nabla_nt)$. Thus

$$\begin{aligned}D_\mu N&=-N^2D_\mu(\nabla_nt)=-N^2q_\mu^\nu n^\rho(\nabla_\rho\nabla_\nu t)\\ &=-sN(\nabla_nn_\nu)=-sN\nabla_nn_\mu\end{aligned} \qquad \text{(I.1.1. 54)}$$

where in the first step we interchanged the second derivative due to torsion freeness and could pull n^ρ out of the second derivative because the correction term is proportional to $n_\rho \nabla n^\rho = 0$ and in the second we have pulled in a factor of N, observed that the correction is annihilated by the projection, used once more $sN\nabla t = n$ and finally used that $\nabla_n n_\nu$ is already spatial. The second derivative then gives simply

$$\begin{aligned} D_\mu D_\nu N &= = -s(D_\mu N)\nabla_n n_\nu - sNq_\mu^\rho q_\nu^\sigma \nabla_\rho \nabla_n n_\sigma \\ &= N(\nabla_n n_\mu)(\nabla_n n_\nu) - sNq_\mu^\rho q_\nu^\sigma \nabla_\rho \nabla_n n_\sigma \end{aligned} \tag{I.1.1. 55}$$

which is indeed N times (I.1.1. 53) as claimed. Notice that in (I.1.1. 55) we cannot replace N by $|N|$ if N is not everywhere positive so *the interpretation that we are driving at would not hold if we would not set $N = |N|$ everywhere*. It is at this point that we must take N positive in all that follows.
We have thus established the key result

$$\begin{aligned} \mathcal{L}_{Nn} K_{\mu\nu} &= N(-KK_{\mu\nu} + 2K_{\rho\mu}K_\nu^\rho) - s[D_\mu D_\nu N \\ &\quad + N(R_{\rho\sigma}^{(D+1)} q_\mu^\rho q_\nu^\sigma - R_{\mu\nu}^{(D)})] \end{aligned} \tag{I.1.1. 56}$$

In order to finish the calculation for $\mathcal{L}_{Nn}P^{\mu\nu}$ we need to know $\mathcal{L}_{Nn}\sqrt{\det(q)}, \mathcal{L}_{Nn}q^{\mu\nu}$. So far we have defined $\det(q)$ in the ADM frame only, its generalization to an arbirtrary frame is given by

$$\det((q_{\mu\nu})(X)) := \frac{1}{D!}[(\nabla_{\mu_0} t)(X)\epsilon^{\mu_0..\mu_D}][(\nabla_{\nu_0} t)(X)\epsilon^{\nu_0..\nu_D}]q_{\mu_1\nu_1}(X)..q_{\mu_D\nu_D}(X) \tag{I.1.1. 57}$$

as one can check by specializing to the ADM coordinates $X^\mu = t, x^a$. Here $\epsilon^{\mu_0..\mu_D}$ is the metric independent, totally skew Levi-Civita tensor density of weight one. One can verify that with this definition we have $\det(g) = sN^2 \det(q)$ by simply expanding $g = q + sn \otimes n$. It is important to see that $\mathcal{L}_T \nabla_\mu t = \mathcal{L}_N \nabla_\mu t = 0$ from which then follows immediately that

$$\mathcal{L}_{Nn}\sqrt{\det(q)} = \frac{1}{2}\sqrt{\det(q)}q^{\mu\nu}\mathcal{L}_{Nn}q_{\mu\nu} = N\sqrt{\det(q)}K \tag{I.1.1. 58}$$

where (I.1.1. 6) has been used. Finally, using once more (I.1.1. 54) we find indeed

$$\mathcal{L}_{Nn}q^{\mu\nu} = -q^{mu\rho}q^{\nu\sigma}\mathcal{L}_{Nn}q_{\rho\sigma} = -2NK^{\mu\nu} \tag{I.1.1. 59}$$

We are now in position to compute the Lie derivative of $P^{\mu\nu} = -s\sqrt{\det(q)}[q^{mu\rho}q^{\nu\sigma} - q^{mu\nu}q^{\rho\sigma}]K_{\rho\sigma}$. Putting all six contributions carefully together and comparing with (I.1.1. 49) one finds the non-trivial result

$$\begin{aligned} \{H(N), P^{\mu\nu}\} &= \frac{q^{\mu\nu}NH}{2} - N\sqrt{\det(q)}[q^{mu\rho}q^{\nu\sigma} - q^{mu\nu}q^{\rho\sigma}]R_{\rho\sigma}^{(D+1)} \\ &\quad + \mathcal{L}_{Nn}P^{\mu\nu} \end{aligned} \tag{I.1.1. 60}$$

that is, only on the constraint surface and only when the (vacuum) equations of motion hold, can the Hamiltonian flow of $P^{\mu\nu}$ with respect to $H(N)$ be interpreted as the action of a diffeomorphism in the direction perpendicular to Σ_t. Now, using again the definition of curvature as the commutator of covariant derivatives it is not difficult to check that

$$\begin{aligned} G_{\mu\nu}n^\mu n^\nu &= \frac{sH}{2\sqrt{\det(q)}} \\ G_{\mu\nu}n^\mu q_\rho^\nu &= -\frac{sH_\rho}{2\sqrt{\det(q)}} \end{aligned} \tag{I.1.1. 61}$$

so that the constraint equations actually are equivalent to $D+1$ of the Einstein equations. Since (I.1.1. 60) contains besides H all the spatial projections of $G_{\mu\nu}$ we see that our interpretation of $\{H(N), P^{\mu\nu}\}$ holds only on shell, $G_{\mu\nu}=0$.
This finishes our geometric analysis of the Hamiltonian flow of the constraints which shows that the symmetry group of spacetime diffeomorphisms Diff(M) of Einstein's action is faithfully implemented in the canonical framework, although in a not very manifest way (more precisely, it is only the subset of those symmetries [50] generated by the Lie algebra of that symmetry group). The importance of this result cannot be stressed enough: It is often said that every $(D+1)-$ diffeomorphism invariant quantity should be a Dirac observable since Diff(M) is the symmetry of the Einstein-Hilbert action. But this would mean that any higher derivative theory (containing arbitrary scalars built from polynomials of the curvature tensor) would also have the same Dirac observables, meaning that to be an observable would be theory independent. The catch is that $(D+1)$dimensional diffeomorphism invariance is not only a kinematical statement but involves the theory dependent dynamics. The fact that the motions generated by the constraints can be interpreted as spacetime diffeomorphisms only on (the theory dependent) shell spells this out in the precise way.

What do these considerations tell us ? The Hamiltonian of general relativity is not a true Hamiltonian but a linear combination of constraints. Rather than generating time translations it generates spacetime diffeomorphisms. Since the parameters of these diffeomorphisms, N, N^a are completely arbitrary unspecified functions, the corresponding motions on the phase space have to be interpreted as *gauge transformations*. This is quite similar to the gauge motions generated by the Gauss constraint in Maxwell theory [24]. The basic variables of the theory, q_{ab}, P^{ab} are not observables of the theory because they are not gauge invariant. Let us count the number of kinematical and dynamical (true) degrees of freedom : The basic variables are both symmetric tensors of rank two and thus have $D(D+1)/2$ independent components per spatial point. There are $D+1$ independent constraints so that $D+1$ of these phase space variables can be eliminated. $D+1$ of the remaining degrees of freedom can be gauged away by a gauge transformation leaving us with $D(D+1)-2(D+1)=(D-2)(D+1)$ phase space degrees of freedom or $(D-2)(D+1)/2$ configuration space degrees of freedom per spatial point. For $D=3$ we thus recover the two graviton degrees of freedom.

The further classical analysis of this system could now proceed as follows :
1) One determines a complete set of gauge invariant observables on the constraint surface $\overline{\mathcal{M}}$ and computes the induced symplectic structure $\overline{\Omega}$ on the so reduced symplectic manifold $\hat{\mathcal{M}}$. Equivalently, one obtains the full set of solutions to the equations of motion, the set of Cauchy data are then the searched for observables. This programme of "symplectic reduction" could never be completed due to the complicated appearance of the Hamiltonian constraint. In fact, until today one does not know any observable for full general relativity (with exception of the generators of the Poincaré group at spatial infinity in the case that (σ, q_{ab}) is asymptotically flat [48]).
2) One fixes a gauge and solves the constraints. Years of research in the field of solving the Cauchy problem for general relativity reveal that such a procedure works at most locally, that is, there do not exist, in general, global gauge conditions. This is reminiscent of the Gribov problem in non-Abelian Yang-Mills theories.

In summary, general relativity can be cast into Hamiltonian form, however, its equations of motion are complicated non-linear partial differential equations of second order and very difficult to solve. Nevertheless, the Cauchy problem is well-posed and the classical theory is consistent up to the point where singularities (e.g. black holes) appear [47]. This is one instance where it is expected that the classical theory is unable to describe the system appropriately any longer and that the more exact theory of quantum gravity must take over in order to remove the singularity. This is expected to be quite in analogy to the case of the hydrogenium atom whose stability was a miracle to classical electrodynamics but was easily explained by quantum physics. Of course, the quantum theory of gravity is expected to be even harder to handle mathematically than the classical theory, however, as a zeroth step an existence proof would already be a triumph. Notice that up to date a similar existence proof for, say, QCD is lacking as well [32].

I.1.2 The Programme of Canonical Quantization

In this section we briefly summarize which steps the method of canonical quantization consists of. We will not go too much into details, the interested reader is referred to [53, 54, 55, 56]. We then consider the application of this programme to canonical general relativity in the ADM formulation and point out the immediate problems that one is confronted with and which prevented one to make progress in this field for so long.

I.1.2.1 The General Programme

Let be given an (infinite dimensional) constrained symplectic manifold $(\mathcal{M}, \Omega)$ modelled on a Banach space E with strong symplectic structure Ω and first class constraint functionals $C_I(N^I)$ (in case of second class constraints one should replace Ω by the corresponding Dirac bracket [24]; there could also be an additional true Hamiltonian which is not constrained to vanish but which is supposed to be gauge invariant). Here I takes values in some finite index set and $C_I(N^I)$ is an appropriate pairing as in the previous section between the constraint density $C_I(x)$, x a point in the D-dimensional manifold of the Hamiltonian framework, and its corresponding Lagrange multiplier N^I. Unless otherwise specified no summation over repeated indices I is assumed.
The quantization algorithm for this system consists of the following.

I) *Polarization*
The phase space can be coordinatized in many ways by what are called "elementary variables", that is, global coordinates such that all functions on $\mathcal{M}$ can be expressed in terms of them. One set of elementary variables may be more convenient than another in the sense that the equations of motion or the constraint functions C_I look more or less complicated in terms of them.

Also, one has to split the set of elementary variables into "configuration $q_a(x)$ and momentum variables $P^a(x)$" (here x is a coordinate of the D-dimensional time slice and a takes values in a finite set). This means that, roughly, wave functions should only depend on half of the number of elementary variables. In the theory of geometric quantization this "splitting" is called a polarization of the symplectic manifold [51]. Among the possible choices of elementary variables those are preferred that come in canonically conjugate pairs $(P^a(x), q_a(x))$, that is, global Darboux coordinates in terms of which the symplectic structure looks as simple as possible. This is important as the quantization of the elementary variables requires that their commutator algebra mirrors their Poisson algebra, see below. In general, the set of elementary variables should form a subalgebra of the Poisson algebra on $\mathcal{M}$ and should be closed under complex conjugation.

Further complications may arise in case that the phase space does not admit an independent set of global coordinates. In this case it may be necessary to work with an overcomplete set of variables and to impose their relations among each other as conditions on states on the Hilbert space. Example :
Suppose we want to coordinatize the cotangent bundle over the sphere S^2. The sphere cannot

be covered by a single coordinate patch, but we can introduce Cartesian coordinates on $\mathbb{R}^3$ and impose the condition $(\hat{x}^1)^2 + (\hat{x}^2)^2 + (\hat{x}^3)^2 - 1 = 0$ on states depending on $\mathbb{R}^3$.

Finally, in the infinite-dimensional context, it is important how to smear the elementary variables : for instance, the relation $\{P(f), P(g)\} = 0$ where $P(f) = \int d^D x f_a(x) P^a(x)$ and f_a is a smooth co-vector field leads one to conclude the distributional relation $\{P^a(x), P^b(y)\} = 0$. However, this is meaningless without specifying the space $\mathcal{S}$ to which the smearing fields f_a belong. For instance, we could also write $\{P^a(x), P^b(y)\} = J^{ab}(x)\delta(x, y)$ where $J^{ab}(x)$ is a non-singular anti-symmetric tensor field supported in a set of $d^D x$ measure zero without affecting $\{P(f), P(g)\} = 0$. If, on the other hand, one would use distributional f^a then $\{P(f), P(g)\}$ may not vanish. Of course, this last point is not independent of the choice of the model space E mentioned above since some model spaces allow distributional smearing fields while others do not.

II) *Quantum Configuration Space*
As experience shows, while the (restriction to the configuration space $\mathcal{C}$ of the) phase space $\mathcal{M}$ is typically some space of smooth fields, complete in a suitable norm, the states of the quantum theory will depend on a more general, distributional quantum configuration space $\overline{\mathcal{C}}$. For instance, the canonical quantization of a free, massive, real scalar field [32] comes with a Hilbert space which is an L_2 space with respect to a Gaussian measure for the corresponding covariance. As is well known, Gaussian measures are supported on a space $\overline{\mathcal{C}}$ of tempered distributions on $\mathbb{R}^n$ and the space $\mathcal{C}$ is contained in a measurable set $\mathcal{N}$ which has measure zero.

Thus, one has to decide what the quantum configuration space $\overline{\mathcal{C}}$ should be. In a sense, this is determined to a large extent by the space $\mathcal{S}$ of the smearing fields F^a : if $q(F)$ is supposed to be a meaningful random variable for a measure (the quantum field $\overline{\mathcal{C}} \ni q : \mathcal{S} \to \mathcal{R}(\mathcal{B});\ F \to F(q)$ is a map from the smearing fields into the random variables of some measureable space $\mathcal{B}$ and is called a generalized stochastic process in the language of probability theory and constructive quantum field theory [32]) then the object q typically lies in the topological dual $\mathcal{S}'$ of $\mathcal{S}$ in view of the Bochner-Minlos theory [32, 58].

III) *Kinematical Measures*
One now has to equip $\overline{\mathcal{C}}$ with the structure of a Hilbert space $\mathcal{H}$. This will be naturally an L_2 space for a suitable measure μ_0 on $\overline{\mathcal{C}}$. Certainly, $\mathcal{H}$ is not yet the physical Hilbert space as one still has to impose the constraints, however, one has to start from such a "kinematical" Hilbert space $\mathcal{H}$ in order to quantize the constraints (the name "kinematical" stems from the fact that it does not know about the constraints yet which capture the *dynamics* of the system). The minimal requirements on a measure μ_0 are as follows :
A)
It should not only be a cylindrical measure but must be $\sigma-$additive, in other words, one must be able to integrate functions of an infinite number of degrees of freedom.
B)
The Hilbert space $\mathcal{H} := L_2(\overline{\mathcal{C}}, d\mu_0)$ must be an irreducible representation of the canonical commutation relations. More precisely, if $F(\hat{q})$ and $\hat{P}(f)$ are the representations of $F(q)$ and $P(f)$ as linear operators on $\mathcal{H}$ with common dense domain $\mathcal{D}$ which they leave invariant, then we must get, for instance, $[\hat{P}(f), F(\hat{q})] = i\hbar\{P(f), F(q)\}^\wedge$. Notice that this condition is well-defined, first because by assumption the Poisson bracket can be expressed in terms of elementary variables again and, secondly the commutator makes sense because by assumption

the elementary operators have domain and range $\mathcal{D}$. Typically, $F(\hat{q})$ will act by multiplication, $(F(\hat{q})\Psi)[q'] = F(q')\Psi[q']$ where the prime is to indicate that q' is distributional rather than smooth while $\hat{P}(f)$ will act as some kind of derivative operator.
Irreducibility of this representation means that the basic operators have a dense range when acting on a cyclic vector. In a reducible representation one tends to have too many degrees of freedom since every irreducible subspace already represents a quantization of the corresponding classical system.

Finally, in case that we have to work with an overcomplete system of elementary variables we must require that the quantizations of their relations among each other are identically satisfied on $\mathcal{H}$.
C)
The Hilbert space must implement the classical complex conjugation relations among the elementary variables as adjointness relations on the corresponding operators. More precisely, by assumption the Poisson subalgebra of the elementary variables is closed under complex conjugation, therefore, e.g., a relation of the kind $\overline{F(q)} = F'_F(q) + P(f'_F)$ will hold for certain smearing fields F'_F, f'_F depending on F. We then require that $F(\hat{q})^\dagger = F'_F(\hat{q}) + \hat{P}(f'_F)$ holds. Here the dagger denotes the adjoint with respect to μ_0. Notice that this condition means also that the domains of the elementary operators and their adjoints coincide.

In summary, we have a representation of the classical * subalgebra of the Poisson algebra, given by the elementary variables, on the Hilbert space.

One can slightly relax these requirements as follows :
Given the classical configuration manifold $\mathcal{C}$ modelled on a Banach space we are naturally equipped also with the space of smooth functions $\mathcal{F}(\mathcal{C})$ and of smooth vector fields $\mathcal{V}(\mathcal{C})$ on it. Let us consider elements (a, b) of the product space $\mathcal{F} \times \mathcal{V}$ and let us equip it with a Lie algebra structure given by $[(a,b),(a',b'))] := (b[a'] - b'[a], [b,b'])$ where $\mathcal{V} \times \mathcal{F} \mapsto \mathcal{F};\ (b,a) \mapsto b[a]$ denotes the natural action of vector fields on functions and $\mathcal{V} \times \mathcal{V} \mapsto \mathcal{V};\ (b,b') \mapsto [b,b']$ denotes the natural action of vector fields on themselves by the Lie bracket, that is, $([b,b'])[a] := b[b'[a]] - b'[b[a]]$. It is easy to see that the set of elementary kinematical variables $F(q), P(f)$ can be identified with points in the set $\mathcal{F} \times \mathcal{V}$ by $F(q) \mapsto (Q_F, 0),\ P(f) \mapsto (0, P_f)$ where $Q_F(q) := F(q),\ (P_f[a])(q) := \{P(f), a\}(q)$. The advantage is now that while the $F(q), P(f)$ may not (be known to) form a closed subalgebra of the Poisson algebra on $\mathcal{M}$, the set $\mathcal{F} \times \mathcal{V}$ always forms a closed Lie algebra. In other words, it may happen that $\{P(f), P(f')\}$ cannot be (obviously) written as a function of the $P(f), F(q)$ again but it is always true that $[P_f, P_{f'}]$ is an element of $\mathcal{V}$ again. It is easy to see that $\{P(f), P(f')\} \mapsto [P_f, P_{f'}]$ if and only if the Jacobi identity $\{\{P(f), P(f'\}, a\} + \text{cyclic} = 0$ holds for all a which, of course, requires the knowledge of $\{P(f), P(f')\}$. Now, the quantization map is given by $(Q_F, P_f) \mapsto (F(\hat{q}), \hat{P}(f))$ and the requrement on the measure is that this be a Lie algebra homorphism. We will take this more general approach also in our case.

IV) *Constraint Operators*
By assumption we can write the classical constraint functions $C_I(N^I)$ as certain functions $C_I(N^I) = c_I(N^I, \{F(q)\}, \{P(f)\})$ of the elementary variables where the curly brackets denote dependence on an in general infinite collection of variables. A naive quantization procedure would be to define its quantization as $\hat{C}_I(N^I) = c_I(N^I, \{F(\hat{q})\}, \{\hat{P}(f)\})$. This will in general not work, at least not straightforwardly, for several reasons :
A)

As is well-known, the quantization of a phase space function is not unique, to a given candidate we can add arbitrary $\hbar$ corrections and still the classical limit of the corrected operator will be the original function. This is called the *factor ordering ambiguity.*
B)
While such corrections in quantum mechanics are relatively harmless, in quantum field theory they tend to be desasterous, a simple example is quantum Maxwell theory where the straightforward quantization of the Hamiltonian gives a divergent nowhere defined operator. It is only after factor ordering that one obtains a densely defined operator. This is what is called a *factor ordering singularity.*
C)
More seriously, in general the singularities of an operator are of an even worse kind and cannot be simply removed by a judicious choice of factor ordering. One has to introduce a regularization of the operator and subtract its divergent piece as one removes the regulator again. This is called the *renormalization* of the operator. The end result must be a densely defined operator on $\mathcal{H}$.
D)
If $C_I(N^I)$ is classically a real-valued function then one would like to implement $C_I(N^I)$ as a self-adjoint operator on $\mathcal{H}$, the reason being that this would guarantee that its spectrum (and therefore its measurement values) is contained in the set of real numbers. While this is certainly a necessary requirement if $C_I(N^I)$ was a true Hamiltonian (i.e. not a constraint), in the case of a constraint this condition can be relaxed as long as the value 0 is contained in its spectrum because this is what we are interested in. On the other hand, a self-adjoint constraint operator is sometimes of advantage when it comes to actually solving the constraints [55, 56].

V) *Imposing the Constraints*
We would now like to solve the constraints in the quantum theory. A first guess of how to do that is by saying that a state $\psi \in \mathcal{H}$ is physical provided that $\hat{C}_I(N^I)\psi = 0$. The study of the simple example of a particle moving in $\mathbb{R}^2$ with the constraint $C = p_2$ reveals that this does not work in general : in the momentum representation $\mathcal{H} = L_2(\overline{\mathcal{C}} := \mathbb{R}^2, d\mu_0 := d^2p)$ the physical state condition becomes $p_2\psi(p_1, p_2) = 0$ with the general solution $\psi_f(p_1, p_2) = \delta(p_2)f(p_1)$ for some function f. The problem is that ψ_f is not an element of $\mathcal{H}$. This is a frequent problem of an operator with continuous spectrum : such an operator does not in general have eigenfunctions in the ordinary sense. However, it has so-called "generalized eigenfunctions" of which ψ_f is an example [59].
The way to solve the constraint is as follows (see [59] for details and section III.7) : One takes a convex topological vector space $\mathcal{D}$ which is dense in $\mathcal{H}$ in the topology of $\mathcal{H}$ and which serves as a common domain for constraint operators and elementary operators. It is then true that $\mathcal{H}$ is contained in the space $\mathcal{D}^*$ of all linear functionals on $\mathcal{D}$ (i.e. $\mathcal{D}^*$ is the algebraic dual of $\mathcal{D}$). We thus have $\mathcal{D} \subset \mathcal{H} \subset \mathcal{D}^*$ (in case that the topology of $\mathcal{D}$ is nuclear and we take the topological dual $\mathcal{D}'$ instead of the algebraic dual, this triple of spaces is called a Gel'fand triple). We now say that an element Ψ of $\mathcal{D}^*$ is a solution of the constraints iff

$$\Psi(\hat{C}_I(N^I)\psi) = 0 \qquad \text{(I.1.2. 1)}$$

for all $I, N^I \in \mathcal{S}, \psi \in \mathcal{D}$.
In the example above, we could take for $\mathcal{D}$ the space of functions of rapid decrease on $\mathbb{R}^2$ and then $\mathcal{D}'$ as the space of tempered distributions on $\mathbb{R}^2$.
The set of solutions in $\mathcal{D}^*$ is called $\mathcal{D}^*_{phys}$. Notice that $\mathcal{D}^*_{phys}$ does not carry a natural Hilbert space structure yet. In the example it is of course natural to take $<\Psi_f, \Psi_g>_{phys} := \int dp_1 \bar{f} g$.

VI) *Quantum Anomalies*
Even if we finally managed to produce a densely defined, possibly self-adjoint operator, with a non-trivial kernel in the above sense we might encounter a quantum anomaly of the following kind :
Recall that by assumption the constraint algebra is first class. This means that there exist so-called structure maps $f_{IJ}{}^K$ from $\mathcal{S}^2$ into the functions on $\mathcal{M}$ such that $\{C_I(N^I), C_J(N^J)\} = \sum_K C_K(f_{IJ}{}^K(N^I, N^J))$. The quantum version of this condition is

$$[\hat{C}_I(N^I), \hat{C}_J(N^J)]\psi = \sum_K [C_K(f_{IJ}{}^K(N^I, N^J))]^\wedge \psi \qquad \text{(I.1.2. 2)}$$

for all $\psi \in \mathcal{D}$. There are two potential problems with (I.1.2. 2) :
First of all, it does not make sense to take a commutator unless the range of the first operator is contained in the domain of the second. Therefore, we must require that all operators $\hat{C}_I(N^I)$ leave $\mathcal{D}$ invariant.
Secondly, notice that especially since, as it is the case in general relativity (see I.1.1. 37), $f_{IJ}{}^K$ depends in general on the phase space coordinates, we are not guaranteed that the right hand side of (I.1.2. 2) can actually be written in the form $\sum_K \hat{C}_K(\hat{f}_{IJ}{}^K(N^I, N^J))\psi$ with the $\hat{C}_K$ *ordered to the left*. If that is not the case then the following inconsistency arises : Let $\Psi \in \mathcal{D}^*_{phys}$ and let us evaluate Ψ on (I.1.2. 2). Then we find that

$$0 = \sum_K \Psi([C_K(f_{IJ}{}^K(N^I, N^J))]^\wedge \psi) \qquad \text{(I.1.2. 3)}$$

for all $\psi \in \mathcal{D}, I, J, N^I, N^J$. Thus, not only does every member of $\mathcal{D}^*_{phys}$ satisfy the constraints (I.1.2. 1) but also the additional constraints (I.1.2. 3) which are absent in the classical theory. Since (I.1.2. 3) will in general be new constraints, algebraically independent from the original ones, the number of physical degrees of freedom in the classical and the quantum theory would differ from each other.
In summary, we must make sure that (I.1.2. 3) is automatically satisfied once (I.1.2. 1) holds which puts additional restrictions on the freedom to order the constraint operators if at all possible.

VII) *Physical Scalar Product*
Suppose that we managed to produce densely defined, anomaly-free constraint operators and the space of solutions $\mathcal{D}^*_{phys}$. How can we arrive at a Hilbert space $\mathcal{H}_{phys}$ with respect to which which the solutions are square integrable ? In the most general case not much is known about a rigorous solution to this problem but an idea is provided by the following formal ansatz in fortunate cases :
Suppose that the constraint operators are all self-adjoint and mutually commuting, then we can formally define the functional δ-distribution

$$\delta[\hat{C}] := \lim_{\epsilon \to 0} \prod_{I, \alpha \in A_\epsilon} \delta(\hat{C}_I(\chi_\alpha)) \qquad \text{(I.1.2. 4)}$$

through the spectral theorem.
The χ_α are the characteristic functions of mutually non-overlapping regions B_α in σ of coordinate volume ϵ^D and $\cup_{\alpha \in A_\epsilon} B_\alpha = \sigma$. Given an element $f \in \mathcal{D}$ we define an element of $\mathcal{D}^*_{phys}$ and a physical inner product between such elements by

$$\Psi_f := \delta[\hat{C}] \cdot f \text{ and } <\Psi_f, \Psi_g>_{phys} := <\delta[\hat{C}] \cdot f, g> \qquad \text{(I.1.2. 5)}$$

where $< .,. >$ is the inner product of $\mathcal{H}$. The fact that the constraints are Abelian reveals that the inner product (I.1.2. 5) is formally Hermitean, non-negative and sesquilinear. The completion of $\mathcal{D}^*_{phys}$ (possibly after factoring by a subspace of null vectors) with respect to $< .,. >_{phys}$ defines the physical Hilbert space $\mathcal{H}_{phys}$.
For a successful application of these ideas in the finite dimensional context see [55, 56] and in the infinite-dimensional and even non-Abelian context see [54, 60].

VIII) *Observables*
By definition, a quantum observable is a self-adjoint operator on $\mathcal{H}_{phys}$. It is actually not difficult to construct such abstractly defined observables once $\mathcal{H}_{phys}$ is known, however, the real problem is to find observables which are quantizations of classical observables, the latter being gauge invariant functions on the constraint surface of the phase space. Obviously, this is a hard problem if not even the classical observables are known as it is the case with pure general relativity. The situation improves if one couples matter [61].
In any case, the physics of the system lies in studying the spectra of the observables. This will in general be a hard problem as well and approximation methods have to be used. General relativity poses a further problem : since there is no true Hamiltonian, the physical Hilbert space is a space "without dynamics". This is the problem of time. There are literally hundreds of publications on this issue without clear conclusions and nothing will be said about it in this article and the author will not even try to give references. However, the author agrees with Rovelli (see [62] and references therein) that the evolution of one physical quantity in a theory without background metric and thus no background time can only be studied relative to another one. Therefore, it is possible to assign to one of the degrees of freedom, say $\hat{O}_1$, the role of a clock variable and one may ask the question : "What is the expectation value of $\hat{O}_2$ in the state Ψ when $\hat{O}_1$ has the expectation value t in the state Ψ ?

This is the outline of the general programme. We will now estimate how far one can get with this programme in application to canonical general relativity as displayed in the previous subsection.

I.1.2.2 Application to General Relativity in the ADM Formulation

Let us go through the steps of the programme one by one and see what the immediate problems are :

I) *Polarization*
Let us assume, as it is usually done throughout the literature, that we choose as elementary variables the ones of the previous section, $F(q) := \int_\sigma d^3x q_{ab} F^{ab}, P(f) := \int_\sigma d^3x P^{ab} f_{ab}$. All fields are smooth and are symmetric tensors with the appropriate density weight. We choose the polarization that the configuration variables be the $F(q)$.

II) *Quantum Configuration Space*
Experience from scalar quantum field theory motivates to have q_{ab} take values in the space of tempered distributions on σ.

III) *Kinematical Measures*
Notice that $\overline{\mathcal{C}}$ is an infinite dimensional noncompact space. Therefore [58] we cannot take μ_0 to be an infinite product Lebesgue measure but must take some sort of probability measure, for instance a Gaussian measure with "white noise" covariance $C_{ab,cd}(x,y) = \frac{1}{\sqrt{\det(q^0)(x)}} q^0_{ac}(x) q^0_{bd}(x) \delta(x,y)$

where q^0 is any fixed positive definite background metric on σ. In other words, the characteristic functional [32] of this measure is given by

$$\chi(f)=\int_{\overline{\mathcal{C}}} d\mu(q)e^{iF(q)}=\exp(-\frac{1}{2}\int d^3x\frac{1}{\sqrt{\det(q^0)}}F^{ab}F^{cd}q^0_{ac}q^0_{bd}) \ . \qquad \text{(I.1.2. 6)}$$

From the general theory [32] we know that this measure is supported on $\overline{\mathcal{C}}$ and that finite linear combinations of states of the form $(\psi_F)(q):=\exp(iF(q))$ are dense in $\mathcal{H}$.

It is obvious that this measure fails to be invariant under three-dimensional diffeomorphisms which turns out to be a major obstacle in solving the diffeomorphism constraint since, for instance, the natural representation of the spatial diffeomorphism group Diff(σ) on $\mathcal{H}$, densely defined by $\hat{U}(\varphi)\psi_F=\psi_{\varphi^*F}$, is not unitary and therefore cannot be generated by a self-adjoint constraint operator ($\varphi\in$Diff(σ) and φ^* is the pull-back action).
In fact, the author is not aware of any work where a diffeomorphism invariant measure was rigorously defined for the stochastic process corresponding to metric quantum fields.

If we let $F(\hat{q})$ act by multiplication and define

$$\hat{P}(f)\psi_F:=\hbar[F(f)+\frac{i}{2}\int d^3x\frac{f_{ab}}{\sqrt{\det(q_0)}}q_0^{ac}q_0^{bd}q_{cd}]\psi_F \qquad \text{(I.1.2. 7)}$$

then the canonical commutation relations and the adjointness relations are, at least formally, indeed satisfied.

IV) *Constraint Operators*
So far we did not encounter any particular problems. However, now we will encounter a major roadblock : Looking at the algebraic structure of (I.1.1. 29) we see that the classical constraint functions depend non-polynomially, not even analytically on the metric q_{ab} (recall that the curvature scalar depends also on the inverse metric tensor q^{ab}). This, first of all, seems to rule out completely the polarization for which the $\hat{P}(f)$ are diagonal since then the $F(\hat{q})$ would become derivative operators. More seriously, since the $\hat{P}^{ab}(x),\hat{q}_{ab}(x)$ are operator valued distributions which are multiplied at the same point in (I.1.1. 29), a simple replacement of variables by operators is hopelessly divergent and completely meaningless since it is not clear how a distribution in the denominator can be defined.

The only chance is that one can suitably regularize expressions (I.1.1. 29) by defining them as limits of functions of smeared field variables, the limit corresponding to vanishing smearing volume. However, nobody succeeded up to date to accomplish such a regularization and renormalization procedure for the quantum operator corresponding to the Hamiltonian constraint (I.1.1. 29) which is also famously called the *Wheeler-DeWitt Operator*.

Since the subsequent steps of the quantization programme depend on this one which we could not solve, not much can be said about the remaining steps.

V) *Imposing the Constraints*
Of course, one could try to find formal solutions to the Wheeler-DeWitt constraint equation which can be seen as the *Quantum-Einstein-Equations*. Not even one solution could be found in the full theory (although solutions could be found in certain finite-dimensional truncations of the theory). Notice that not even the constant state $\psi(q)=1$ is a solution.

VI) *Quantum Anomalies*
Given that one could not even define the Wheeler-DeWitt constraint operator it seems to be a hopeless enterprise to find an ordering for which it is free of anomalies or even self-adjoint.

VII) *Physical Scalar Product*
DeWitt has defined in his famous three works [23] a formal inner product which is at least invariant under three-dimensional diffeomorphisms, however, to the best knowledge of the author nobody could ever give a rigorous meaning to the construction.

VIII) *Observables*
This step has been completely out of reach since one started the analysis.

Summarizing, the programme of canonical quantization applied the way as just displayed was unsuccessful for decades. Thus, most researchers in the field gave up and turned to different approaches. It should be kept in mind, however, that the programme is not a rigid algorithm but requires to make choices at various stages which are not dictated by mathematical consistency but depend on one's intuition. Already in the very first step one is asked to make a choice about the elementary variables and the polarization of the phase space. Until the mid 80's people worked only with those ADM variables displayed above since they are so natural. On the other hand, given the complicated structure of (I.1.1. 29) which was a roadblock for such a long time, it seems mandatory to look for better suited canonical variables which, preferrably, render the constraints at least polynomial. This is precisely the achievement of Ashtekar [30].

I.1.3 The New Canonical Variables of Ashtekar for General Relativity

In this subsection we focus on the classical aspects of the so-called "new variables' '. The history of the the classical aspects of the new variables is approximately twenty years old and we wish to give a brief account of the developments (the history of the quantum aspects will be given in section I.1.4):

- *1981-82*
The starting point was a series of papers due to Sen [63] who generalized the covariant derivative ∇_μ of the previous section for $s = -1$ to $Sl(2,\mathbb{C})$ spinors of left (right) handed helicity resulting in an (anti) self-Hodge-dual connection which is therefore complex-valued. An exhaustive treatment on spinors and spinor calculus can be found in [64].

- *1986-87*
Sen was motivated in part by a spinorial proof of the positivity of energy theorem of general relativity [65, 66]. But it was only Ashtekar [30, 45, 53] who realized that modulo a slight modification of his connection, Sen had stumbled on a new canonical formulation of general relativity in terms of the (spatial projection of) this connection, which turns out to be a generalization of D_μ to this class of spinors, and a conjugate electrical field kind of variable, such that the initial value constraints of general relativity (I.1.1. 29) can be written in *polynomial form* if one rescales H by $H \mapsto \tilde{H} = \sqrt{\det(q)}H$ (which looks like a harmless modification at first sight). In fact $\tilde{H}$ is only of fourth order in the canonical coordinates, not worse than non-Abelian Yang-Mills theory and thus a major roadblock on the way towards quantization seemed to be removed. Ashtekar also noted the usefulness of the connection for $s = +1$ in which case it is actually real-valued [67].

- *1987-88*
  Ashtekar's proofs were in a Hamiltonian context. Samuel as well as Jacobson and Smolin dicovered independently that there exists in fact a Lagrangean formulation of the theory by considering only the (anti) selfdual part of the curvature of Palatini gravity [68]. Jacobson also considered the coupling of fermionic matter [69] and an extension to supergravity [70]. Coupling to standard model matter was considered by Ashtekar et. al. in [71]. All of these developments still used a spinorial language which, although not mandatory, is of course quite natural if one wants to treat spinorial matter.

  A purely tensorial approach to the new variables was given by Goldberg [72] in terms of triads and by Henneax et. al. in terms of tetrads [73].

- *1989-92*
  While the Palatini formulation of general relativity uses a connection and a tetrad field as independent variables, Capovilla, Dell and Jacobson realized that there is a classically equivalent action which depends only on a connection and a scalar field, moreover, they were able to solve both initial value constraints of general relativity *algebraically* for a huge (but not the complete) class of field configurations. Unfortunately, there is a third constraint besides the diffeomorphism and Hamiltonian constraint in this new formulation of general relativity, the so-called Gauss constraint, which is not automatically satisfied by this so-called "CDJ-Ansatz' '.

  This line of thought was further developed by Bengtsson and Peldan [74] culminating in the discovery that in the presence of a cosmological constant the just mentioned scalar field can be eliminated by a field equation, resulting in a pure connection Lagrangean for general relativity (but not a polynomial one).

- *1994-96*
  As mentioned above, for Lorentzian (Euclidean) signature one considered complex (real) valued connection variables. Meanwhile it turned out that it is very hard to implement the reality conditions for the complex valued case as adjointness conditions on the measure in the quantum theory while for the real valued case it is relatively easy. This motivated Barbero [76] to consider real valued connections also for Lorentzian signature. Barbero discovered that one can give a Hamiltonian formulation even *for all complex values* of a parameter considered earlier by Immirzi [77] for either choice of signature. However, in order to keep polynomiality of the Hamiltonian constraint when using real valued connections one has to multiply it by an even higher power of $\det(q)$. Moreover, the constraint becomes algebraically much more complicated.

  This caveat is removed by a so-called "phase space Wick rotation' ' intoduced in [80, 81] and later considered also in [82] where one can work with real connections *while keeping the algebraic form of the constraint simple.* This line of development was motivated by a seminal paper due to Hall [79] who constructed a unitary transform from a Hilbert space of square integrable functions on a compact gauge group to a Hilbert space of square integrable, holomorphic functions on the complexification of that gauge group and this transform was generalized by Ashtekar, Lewandowski, Marolf, Mourão and Thiemann to gauge theories for compact gauge groups [83]. Mena-Marugan clarified the relation between this phase space Wick rotation and the usual one (analytic continuation in the time parameter) [84].

  The last development in this respect is the result of [78] which states that polynomiality of the constraint operator is not only unimportant in order to give a rigorous meaning to it in quantum theory, it is in fact *desastrous*. The important condition is that the constraint be a scalar of *density of weight one*. This forbids the rescaling from H, which is already a density of weight one, by any non-trivial power of $\det(q)$. It is only in that case that the quantization of the operator can be done in a *background independent way without picking up UV divergences on the kinematical Hilbert space*. For this reason, real connection variables are currently favoured as far as quantum theory is concerned. In retrospect, what is really important is that one bases the quantum theory on connections and canonically conjugate electric fields (which is dual in a metric independent way to a two-form). The reason is that n-forms can be naturally integrated over n-dimensional submanifolds of σ without requiring a background structure, this is not possible for the metric variables of the ADM formulation and has forbidden progress for such a long time. We will come back to this point in the next section.

- *1996-2000*

  So far a Lagrangean action principle had been given only for the following values of signature s and Immirzi parameter β, namely Lorentzian general relativity $s=-1, \beta=\pm i$ and Euclidean general relativity $s=+1, \beta=\pm 1$. For arbitrary complex β and either signature a Lagrangean formulation was discovered by Holst and Barros e Sá [85]. Roughly speaking the action is given by a modification of the Palatini action

  $$S=\int_M \operatorname{tr}(F\wedge[*-\beta^{-1}](e\wedge e)) \qquad \text{(I.1.3. 1)}$$

  (it results for $\beta=\infty$) where $*$ denotes the Hodge dual, $F=F(\omega)$ is the curvature of some connection ω which is considered as an independent field next to the tetrad e. This action should be considered in analogy with the θ theta angle modification of bosonic QCD

  $$S=\int_M \operatorname{tr}(F\wedge[*+\theta]F) \qquad \text{(I.1.3. 2)}$$

  In the gravitational case the β term drops out by an equation of motion, in the QCD case the variation of the θ term is exact and also drops out of the equations of motion. This holds for the classical theory, but it is well known that in the quantum theory the actions with different values of θ are not unitarily equivalent. A similar result holds for general relativity [77].

  Recently Samuel [86] criticized the use of real connection variables for Lorentzian gravity because of the following reason: The Hamiltonian analysis of the action (I.1.3. 1) leads, unless $\beta=\pm i$ for $s=-1$, to constraints of second class which one has to solve by imposing a gauge condition. It eliminates the boost part of the original $SO(1,3)$ Gauss constraint and one is left with an $SO(3)$ Gauss constraint (which also appears in in the case $\beta=\pm i$). That gauge condition fixes the direction of an internal $SO(1,3)$ vector which is automatically preserved by the remaining $SO(3)$ subgroup and by the evolution derived from the associated Dirac bracket, so that everything is consistent. Now while for $\beta=\pm i, s=-1$ the spatial connection is simply the pull-back of the (anti)self-dual part of the four-dimensional spin connection to the spatial slice, for real β its spacetime interpretation is veiled due to the appearance of the second class constaints and the gauge fixing.

  Samuel now asks the following question: For any value of β it can be shown that every $SO(3)$ gauge invariant function of the spatial connection and the triad can be expressed in terms of the (pull-back to the spatial slice of the) spacetime fields $q_{\mu\nu}, K_{\mu\nu}$. In the previous section we have shown that the Hamiltonian evolution of these fields under the Hamiltonian constraint coincides, on the constraint surface, with their infinitesimal transformation under a timelike diffeomorphism. Is it then true that the induced Hailtonian transformation of $SO(3)$ gauge invariant functions of the connection (such as traces of its holonomy around a loop in a spatial slice) coincides with that of (the pull-back to the spatial slice of) a spacetime connection? He finds that this is the case if and only if $\beta = \pm i$. The simple algebraic reason is that only for an (anti)self-dual connection A^{IJ}, $I, J = 0, 1, 2, 3$ the components A^{0j} are already determined by $A^j = \frac{1}{2}\epsilon_{jkl}A^{kl}$ so that the pull-back to the spatial slice of A^j determines the pull-back of an $SO(1,3)$ connection with its full spacetime interpretation only then.

It should be stressed, however, that Samuel's criticism is purely aesthetical in nature, for interpretational reasons it is certainly convenient to have a spacetime interpretation of the spatial connection but it is by no means mandatory, one just has to bear in mind that the connection does not have the naive transformation behaviour under Hamiltonian evolution on the constraint surface. In fact, to date a satisfactory quantum theory has been constructed only for β real (which in turn does not mean that it is impossible to do for $\beta = \pm i$). In fact, as we will show in this subsection, at the classical level *all complex values of the Immirzi parameter lead to Hamiltonian formulations completely equivalent to the ADM formulation.*

This concludes our historical digression and we come now to the actual derivation of the new variable formulation. We decided for the extended phase space approach using triads as this makes the contact and equivalence with the ADM formulation most transparent and quickest and avoids the introduction of additional $SL(2,\mathbb{C})$ spinor calculus which would blow up our exposition unnecessarily. Also we do this for either signature and any complex value of the Immirzi parameter. What is no longer arbitrary is the dimension of σ : We will be forced to work with $D = 3$ as will become clear in the course of the derivation.
The construction consists of two steps : First an extension of the ADM phase space and secondly a canonical transformation on the extended phase space.

Extension of the ADM phase space

We would like to consider the phase space described in section I.1.1 as the symplectic reduction of a larger symplectic manifold with coisotropic constraint surface [51]. One defines a so-called co-D-bein field e_a^i on σ where the indices $i, j, k, ..$ take values $1, 2, .., D$. The D-metric is expressed in terms of e_a^i as

$$q_{ab} := \delta_{jk} e_a^j e_b^k \ . \tag{I.1.3. 3}$$

Notice that this relation is invariant under local $SO(D)$ rotations $e_a^i \to O_j^i e_a^j$ and we therefore can view e_a^i, for $D = 3$, as an $su(2)$-valued one-form (recall that the adjoint representation of $SU(2)$ on its Lie algebra is isomorphic with the defining representation of $SO(3)$ on $\mathbb{R}^3$ under the isomorphism $\mathbb{R}^3 \to su(2)$; $v^i \to v^i\tau_i$ where τ_i is a basis of $su(2)$ (also called "soldering forms" [93]). This observation makes it already obvious that we have to get rid of the $D(D-1)/2$ rotational degrees of freedom sitting in e_a^i but not in q_{ab}. Since the Cartan-Killing metric of $so(D)$ is just the Euclidean one we will in the sequel drop the δ_{ij} and also do not need to care about index positions.

Next we introduce yet another, independent one form K_a^i on σ which for $D = 3$ we also consider as $su(2)$ valued and from which the extrinsic curvature is derived as

$$-2sK_{ab} := \mathrm{sgn}(\det((e_a^i)))K_{(a}^i e_{b)}^i \ . \tag{I.1.3. 4}$$

We see immediately that K_a^i cannot be an arbitrary $D \times D$ matrix but must satisfy the constraint

$$G_{ab} := K_{[a}^j e_{b]}^j = 0 \tag{I.1.3. 5}$$

since K_{ab} was a symmetric tensor field. With the help of the quantity

$$E^a_j := \frac{1}{(D-1)!}\epsilon^{aa_1..a_{D-1}}\epsilon_{jj_1..j_{D-1}}e^{j_1}_{a_1}..e^{j_{D-1}}_{a_{D-1}} \tag{I.1.3. 6}$$

one can equivalently write (I.1.3. 6) in the form

$$G_{jk} := K_{a[j}E^a_{k]} = 0 \tag{I.1.3. 7}$$

Consider now the following functions on the extended phase space

$$q_{ab} := E^j_a E^j_b |\det((E^c_l))|^{2/(D-1)},\ P^{ab} := |\det((E^c_l))|^{-2/(D-1)}E^a_k E^d_k K^j_{[d}\delta^b_{c]}E^c_j \tag{I.1.3. 8}$$

where E^j_a is the inverse of E^a_j. It is easy to see that when $G_{jk} = 0$, the functions (I.1.3. 8) precisely reduce to the ADM coordinates. Inserting (I.1.3. 8) into (I.1.1. 29) we can also write the diffeomorphism and Hamiltonian constraint as functions on the extended phase space which one can check to be explicitly given by

$$\begin{aligned} H_a &:= -D_b[K^j_a E^b_j - \delta^b_a K^j_c E^c_j] \\ H &:= -\frac{s}{4\sqrt{\det(q)}}(K^l_a K^j_b - K^j_a K^l_b)E^a_j E^b_l - \sqrt{\det(q)}R \end{aligned} \tag{I.1.3. 9}$$

where $\sqrt{\det(q)} := |\det((E^a_j))|^{1/(D-1)}$ and $q^{ab} = E^a_j E^b_j/\det(q)$ by which $R = R(q)$ is considered as a function of E^a_j. Notice that, using (I.1.3. 4), (I.1.3. 6), expressions (I.1.3. 9) indeed reduce to (I.1.1. 29) up to terms proportional to G_{jk}.

Let us equip the extended phase space coordinatized by (K^i_a, E^a_i) with the symplectic structure (formally, that is without smearing) defined by

$$\{E^a_j(x), E^b_k(y)\} = \{K^j_a(x), K^k_b(y)\} = 0,\ \{E^a_i(x), K^j_b(y)\} = \kappa\delta^a_b\delta^j_i\delta(x,y) \tag{I.1.3. 10}$$

We claim now that the symplectic reduction with respect to the constraint G_{jk} of the constrained Hamiltonian system subject to the constraints (I.1.3. 7), (I.1.3. 8) results precisely in the ADM phase space of section I.1.1 together with the original diffeomorphism and Hamiltonian constraint.

To prove this statement we first of all define the smeared "rotation constraints"

$$G(\Lambda) := \int_\sigma d^D x \Lambda^{jk} K_{aj} E^a_k \tag{I.1.3. 11}$$

where $\Lambda^T = -\Lambda$ is an arbitrary antisymmetric matrix, that is, an $so(D)$ valued scalar on σ. They satisfy the Poisson algebra, using (I.1.3. 10)

$$\{G(\Lambda), G(\Lambda')\} = G([\Lambda, \Lambda']) \tag{I.1.3. 12}$$

in other words, $G(\Lambda)$ generates infinitesimal $SO(D)$ rotations as expected. Since the functions (I.1.3. 8) are manifestly $SO(D)$ invariant by inspection they Poisson commute with $G(\Lambda)$, that is, they comprise a complete set of rotational invariant Dirac observables with respect to $G(\Lambda)$ for any Λ. As the constraints defined in (I.1.3. 9) are in turn functions of these, $G(\Lambda)$ also Poisson commutes with the constraints (I.1.3. 9) whence the total system of constraints consisting of (I.1.3. 11), (I.1.3. 9) is of first class.

Finally we must check that Poisson brackets among the q_{ab}, P^{cd}, considered as the functions (I.1.3. 8) on the extended phase space with symplectic structure (I.1.3. 10), is equal to the Poisson brackets of the ADM phase space (I.1.1. 30, at least when $G_{jk}=0$. Since q_{ab} is a function of E^a_j only it is clear that $\{q_{ab}(x),q_{cd}(y)\}=0$. Next we have

$$\begin{aligned}
\kappa\{P^{ab}(x),q_{cd}(y)\} &= (\frac{1}{2}[q^{a(e}q^{bf)}-q^{ab}q^{ef}]E^j_f)(x)\{K^j_e(x),(|\det(E)|^{2/(D-1)}E^k_cE^k_d)(y)\} \\
&= (\frac{1}{2}[q^{a(e}q^{bf)}-q^{ab}q^{ef}]E^j_f)(x)[\frac{2}{D-1}q_{cd}(x)\frac{\{K^j_e(x),|\det(E)|(y)\}}{|\det(E)|(x)} \\
&\quad +2(\det(q)E^k_{(c}(x)\{K^j_e(x),E^k_{d)}(y)\}] \\
&= ([q^{a(e}q^{bf)}-q^{ab}q^{ef}][-\frac{1}{D-1}q_{cd}q_{ef}+q_{e(c}q_{d)f}])(x)\delta(x,y) \\
&= \delta^a_{(c}\delta^b_{d)}\delta(x,y) \qquad\qquad \text{(I.1.3. 13)}
\end{aligned}$$

where we used $\delta E^{-1}=-E^{-1}\delta EE^{-1}$, $[\delta|\det(E)|]/|\det(E)|=[\delta\det(E)]/\det(E)=E^j_a\delta E^a_j$. The final Poisson bracket is the most difficult one. By carefully inserting the definitions, making use of the relations $E^a_j=\det(e)e^a_j, E^j_a=e^j_a\det(e), e^a_j=q^{ab}e^j_b$ at various steps one finds after two pages of simple but tedious algebraic manipulations that

$$\{P^{ab}(x),P^{cd}(y)\}=-\frac{\det(e)}{8}[q^{bc}G^{ad}+q^{bd}G^{ac}+q^{ac}G^{bd}+q^{ad}G^{bc}])(x)\delta(x,y) \qquad \text{(I.1.3. 14)}$$

where $G^{ab}=q^{ac}q^{bd}G_{cd}$ and so (I.1.3. 14) vanishes only at $G_{ab}=0$.

Let us summarize : The functions (I.1.3. 8) and (I.1.3. 9) reduce at $G_{jk}=0$ to the corresponding functions on the ADM phase space, moreover, their Poisson brackets among each other reduce at $G_{jk}=0$ to those of the ADM phase space. Thus, as far as rotationally invariant observables are concerned, the only ones we are interested in, both the ADM system and the extended one are completely equivalent and we can as well work with the latter. This can be compactly described by saying that the symplectic reduction with respect to G_{jk} of the constrained Hamiltonian system described by the action

$$S:=\frac{1}{\kappa}\int_{\mathbb{R}}dt\int_\sigma d^Dx(\dot{K}^j_aE^a_j-[-\Lambda^{jk}G_{jk}+N^aH_a+NH]) \qquad \text{(I.1.3. 15)}$$

is given by the system described by the ADM action of section (I.1.1). Notice that, in accordance with what we said before, there is no claim that the Hamiltonian flow of K^j_a, E^a_j with respect to H_a, H is a spacetime diffeomorphism. However, since the Hamiltonian flow of H, H_a on the constraint surface $G_{jk}=0$ is the same as on the ADM phase space for the gauge invariant observables q_{ab}, P^{ab}, a representation of Diff(M) is still given on the constraint surface of $G_{jk}=0$.

Canonical Transformation on the Extended Phase Space

Up to now we could work with arbitrary $D\geq 2$, however, what follows works only for $D=3$. First we introduce the notion of the *spin connection* which is defined as an extension of the spatial covariant derivative D_a from tensors to generalized tensors with $so(D)$ indices. One defines

$$D_au_{b..}v_j:=(D_au_b)_{..}v_j+..+u_{b..}(D_av_j) \text{ where } D_av_j:=\partial_av_j+\Gamma_{ajk}v^k \qquad \text{(I.1.3. 16)}$$

extends by linearity and requires that D_a is compatible with e^j_a, that is

$$D_ae^j_b=0\ \Rightarrow\ \Gamma_{ajk}=-e^b_k[\partial_ae^j_b-\Gamma^c_{ab}e^j_c] \qquad \text{(I.1.3. 17)}$$

Obviously Γ_a takes values in $so(D)$, that is, (I.1.3. 17) defines an antisymmetric matrix.

Our aim is now to write the constraint G_{jk} in such a form that it becomes the Gauss constraint of an $SO(D)$ gauge theory, that is, we would like to write it in the form $G_{jk} = (\partial_a E^a + [A_a, E^a])_{jk}$ for some $so(D)$ connection A. It is here where $D = 3$ is singled out : What we have is an object of the form E^a_j which transforms in the defining representation of $SO(D)$ while E^a_{jk} transforms in the adjoint representation of $SO(D)$. It is only for $D = 3$ that these two are equivalent. Thus from now on we take $D = 3$.

The canonical transformation that we have in mind consists of two parts : 1) A constant Weyl (rescaling) transformation and 2) an affine transformation.

Constant Weyl Transformation

Observe that for any finite complex number $\beta \neq 0$, called the *Immirzi parameter*, the following rescaling $(K^j_a, E^a_j) \mapsto ({}^{(\beta)}K^j_a := \beta K^j_a, {}^{(\beta)}E^a_j := E^a_j/\beta)$ is a canonical transformation (the Poisson brackets (I.1.3. 10) are obviously invariant under this map). We will use the notation $K = K^{(1)}, E = K^{(1)}$. In particular, for the rotational constraint (which we write in $D = 3$ in the equivalent form)

$$G_j = \epsilon_{jkl} K^k_a E^a_l = \epsilon_{jkl} ({}^{(\beta)}K^k_a)({}^{(\beta)}E^a_l) \qquad \text{(I.1.3. 18)}$$

is invariant under this rescaling transformation. We will consider the other two constraints (I.1.3. 9) in a moment.

Affine Transformation

We notice from (I.1.3. 17) that $D_a E^b_j = 0$. In particular, we have

$$D_a E^a_j = [D_a E^a]_j + \Gamma_{aj}{}^k E^a_k = \partial_a E^a_j + \epsilon_{jkl} \Gamma^k_a E^a_l = 0 \qquad \text{(I.1.3. 19)}$$

where the square bracket in the first identity means that D acts only on tensorial indices which is why we could replace D by ∂ as E^a_j is an $su(2)$ valued vector density of weight one. We also used the isomorphism between antisymmetric tensors of second rank and vectors in Euclidean space to define $\Gamma_a =: \Gamma^l_a T_l$ where $(T_l)_{jk} = \epsilon_{jlk}$ are the generators of $so(3)$ in the defining – or, equivalently, of $su(2)$ in the adjoint representation if the structure constants are chosen to be ϵ_{ijk}. Next we explicitly solve the spin connection in terms of E^a_j from (I.1.3. 17) by using the explicit formula for Γ^a_{bc} and find

$$\begin{aligned} \Gamma^i_a &= \frac{1}{2}\epsilon^{ijk} e^b_k [e^j_{a,b} - e^j_{b,a} + e^c_j e^l_a e^l_{c,b}] \\ &= \frac{1}{2}\epsilon^{ijk} E^b_k [E^j_{a,b} - E^j_{b,a} + E^c_j E^l_a E^l_{c,b}] + \frac{1}{4}\epsilon^{ijk} E^b_k [2E^j_a \frac{(\det(E))_{,b}}{\det(E)} - E^j_b \frac{(\det(E))_{,a}}{\det(E)}] \end{aligned} \qquad \text{(I.1.3. 20)}$$

where in the second line we used that $\det(E) = [\det(e)]^2$ in $D = 3$. Notice that the second line in (I.1.3. 20) explicitly shows that Γ^j_a is a homogenous rational function of degree zero of E^a_j and its derivatives. Therefore we arrive at the important conclusion that

$$({}^{(\beta)}\Gamma^j_a) := \Gamma^j_a({}^{(\beta)}E) = \Gamma^j_a = \Gamma^j_a({}^{(1)}E) \qquad \text{(I.1.3. 21)}$$

is itself invariant under the rescaling transformation. This is obviously also true for the Chritoffel connection Γ^a_{bc} since it is a homogenous rational function of degree zero in q_{ab} and its derivatives and

$q_{ab} = \det(E) E^j_a E^j_b \mapsto ({}^{(\beta)}q_{ab}) = \beta({}^{(1)}q_{ab})$. Thus the derivative D_a is, in fact, independent of β and we therefore have in particular $D_a({}^{(\beta)}E^a_j) = 0$. We can then write the rotational constraint in the form

$$G_j = 0 + \epsilon_{jkl}({}^{(\beta)}K^k_a)({}^{(\beta)}E^a_l) = \partial_a({}^{(\beta)}E^a_j) + \epsilon_{jkl}[\Gamma^j_a + ({}^{(\beta)}K^k_a)]({}^{(\beta)}E^a_l) =: \; {}^{(\beta)}D_a \; {}^{(\beta)}E^a_j \tag{I.1.3. 22}$$

This equation suggests to introduce the new connection

$$({}^{(\beta)}A^j_a) := \Gamma^j_a + ({}^{(\beta)}K^j_a) \tag{I.1.3. 23}$$

This connection could be called the Sen – Ashtekar – Immirzi – Barbero connection (names in historical order) for the historical reasons mentioned in the beginning of this section. More precisely the Sen connection arises for $\beta = \pm i, G_j = 0$, the Ashtekar connection for $\beta = \pm i$, the Immirzi connection for complex β and the Barbero connection for real β. For simplicity we will refer to it as the *new* connection which now replaces the spin-connection Γ^j_a and gives rise to a new derivative ${}^{(\beta)}D_a$ acting on generalized tensors as the extension by linearity of the basic rules ${}^{(\beta)}D_a v_j := \partial_a v_j + \epsilon_{jkl}({}^{(\beta)}A^k_a)v_l$ and ${}^{(\beta)}D_a u_b := D_a u_b$. Notice that (I.1.3. 23) has *precisely* the structure of a Gauss law constraint for an $SU(2)$ gauge theory although ${}^{(\beta)}A$ qualifies as the pull-back to σ by local sections of a connection on an $SU(2)$ fibre bundle over σ only when β is real. Henceforth we will call G_j the *Gauss constraint.*

Given the complicated structure of (I.1.3. 20) it is quite surprising that the variables $({}^{(\beta)}A, {}^{(\beta)}E)$ form a canonically conjugate pair, that is

$$\{{}^{(\beta)}A^j_a(x), {}^{(\beta)}A^k_b(y)\} = \{{}^{(\beta)}E^a_j(x), {}^{(\beta)}E^b_k(y)\} = 0, \; \{{}^{(\beta)}E^a_j(x), {}^{(\beta)}A^j_b(y)\} = \kappa\delta^a_b\delta^k_j\delta(x,y) \tag{I.1.3. 24}$$

This is the key feature for why these variables are at all useful in quantum theory : If we would not have such a simple bracket structure classically then it would be very hard to find Hilbert space representations that turn these Poisson bracket relations into canonical commutation relations.

To prove (I.1.3. 24) by means of (I.1.3. 10) (which is invariant under replacing K, E by ${}^{(\beta)}K, {}^{(\beta)}E$) we notice that the only non-trivial relation is the first one since $\{E^a_j(x), \Gamma^k_b(y)\} = 0$. That relation is explicitly given as

$$\beta[\{\Gamma^j_a(x), K^k_b(y)\} - \{\Gamma^k_b(y), K^j_a(x)\}] = \beta\kappa[\frac{\delta\Gamma^j_a(x)}{\delta E^b_k(y)} - \frac{\delta\Gamma^k_b(y)}{\delta E^a_j(x)}] = 0 \tag{I.1.3. 25}$$

which is just the integrability condition for Γ^j_a to have a generating potential F. A promising candidate for F is given by the functional

$$F = \int_\sigma d^3x E^a_j(x)\Gamma^a_j(x) \tag{I.1.3. 26}$$

since if (I.1.3. 25) holds we have

$$\begin{aligned} \frac{\delta F}{\delta E^a_j(x)} - \Gamma^j_a(x) &= \int d^3y E^b_k(y)\frac{\delta\Gamma^k_b(y)}{\delta E^a_j(x)} = \int d^3y E^b_k(y)\frac{\delta\Gamma^j_a(x)}{\delta E^b_k(y)} \\ = \frac{1}{\kappa}\{\Gamma^j_a(x), \int d^3y K^k_b(y)E^b_k(y)\} = 0 \end{aligned} \tag{I.1.3. 27}$$

because the function $\int d^3y K^k_b(y) E^b_k(y)$ is the canonical generator of *constant* scale transformations under which Γ^j_a is invariant as already remarked above. To show that F is indeed a potential for

Γ^j_a we demonstrate (I.1.3. 27) in the form $\int d^3x E^a_j(x)\delta\Gamma^j_a(x)=0$. Starting from (I.1.3. 20) we have (using $\delta e^j_a e^b_j = \delta e^j_b e^b_k = 0$ repeatedly)

$$\begin{aligned}
e^a_i\delta\Gamma^i_a &= \frac{1}{2}\epsilon^{ijk}\det(e)e^a_i\delta(e^b_k[e^j_{a,b}-e^j_{b,a}+e^c_j e^l_a e^l_{c,b}]) \\
&= \frac{1}{2}\epsilon^{ijk}\det(e)[e^a_i\delta(e^b_k(e^j_{a,b}-e^j_{b,a}))+\delta(e^b_k e^c_j e^i_{c,b})-(\delta e^a_i)e^c_j e^l_a e^b_k e^l_{c,b}] \\
&= \frac{1}{2}\epsilon^{ijk}\det(e)[e^a_i\delta(e^b_k(e^j_{a,b}-e^j_{b,a}))+\delta(e^b_k e^a_j e^i_{a,b})+(\delta e^l_a)e^a_i e^c_j e^b_k e^l_{c,b}] \\
&= \frac{1}{2}\epsilon^{ijk}\det(e)[\delta(e^a_i e^b_k(e^j_{a,b}-e^j_{b,a})+e^b_k e^a_j e^i_{a,b})-(\delta e^a_i)e^b_k(e^j_{a,b}-e^j_{b,a})+(\delta e^l_a)e^a_i e^c_j e^b_k e^l_{c,b}] \\
&= \frac{1}{2}\epsilon^{ijk}\det(e)[\delta(e^b_k(e^a_j e^i_{a,b}+e^a_i e^j_{a,b})-e^a_i e^b_k e^j_{b,a}))+(\delta e^b_k)e^a_i e^j_{b,a}+(\delta e^a_i)e^b_k e^j_{b,a} \\
&\quad +(\delta e^l_a)e^a_i e^c_j e^b_k e^l_{c,b}] \\
&= -\frac{1}{2}\epsilon^{abc}[e^j_c\delta e^j_{b,a}-(\delta e^j_a)e^j_{c,b}] \\
&= -\frac{1}{2}\epsilon^{abc}\partial_a[(\delta e^j_b)e^j_c]
\end{aligned} \tag{I.1.3. 28}$$

From the first to the second line we pulled e^a_i into the variation of the the third term of $\delta\Gamma^a_i$ resulting in a correction proportional to δe^i_a, in the next line we relabelled the summation index c into a in the third term and traded the variation of e^a_i for that of e^l_a in the fourth term, in the next line we pulled again e^a_i inside a variation resulting in altogether six terms, in the next line we collected the total variation terms and reordered them and in the fourth term we relabelled the summation indices a,b into b,a and i,k into k,i resulting in a minus sign from the ϵ^{ijk}, in the next line we realized that the first two terms are symmetric in i,j which thus drop out due to the ϵ^{ijk} and that the e^a_i and e^b_k variation pieces of the third term cancel against the fourth and fifth term, in the next line we made use of the relations $\det(e)\epsilon^{ijk}e^b_j e^c_k=\epsilon^{abc}e^i_a$, $\det(e)\epsilon^{ijk}e^a_i e^b_j e^c_k=\epsilon^{abc}$ and relabelled j for l and in the last line finally we relabelled a for b in the second term resulting in a minus sign and allows us to write the whole thing as a derivative. It follows that

$$\int_\sigma d^3x E^a_j\delta\Gamma^j_a = -\frac{1}{2}\int_\sigma d^3x\partial_a(\epsilon^{abc}\delta e^j_b e^j_c) = \frac{1}{2}\int_{\partial\sigma} dS_a\epsilon^{abc}e^j_b\delta e^j_c \tag{I.1.3. 29}$$

which vanishes since $\partial\sigma=\emptyset$. If σ has a boundary such as spatial infinity then the boundary conditions such as imposing e^j_a to be an even function on the asymptotic sphere under Cartesian coordinate reflection guarantee vanishing of (I.1.3. 29) as well, see [48, 86].

It remains to write the constraints (I.1.3. 9) in terms of the variables ${}^{(\beta)}A, {}^{(\beta)}E$. To that end we introduce the curvatures

$$\begin{aligned}
R^j_{ab} &:= 2\partial_{[a}\Gamma^j_{]a}+\epsilon_{jkl}\Gamma^k_a\Gamma^l_b \\
{}^{(\beta)}F^j_{ab} &:= 2\partial_{[a}\ {}^{(\beta)}A^j_{b]}+\epsilon^{(\beta)}_{jkl}A^k_a\ {}^{(\beta)}A^l_b
\end{aligned} \tag{I.1.3. 30}$$

whose relation with the covariant derivatives is given by $[D_a,D_b]v_j = R_{abjl}v^l = \epsilon_{jkl}R^k_{ab}v^l$ and $[{}^{(\beta)}D_a, {}^{(\beta)}D_b]v_j = {}^{(\beta)}F_{abjl}v^l = \epsilon_{jkl}\ {}^{(\beta)}F^k_{ab}v^l$. Let us expand ${}^{(\beta)}F$ in terms of Γ and ${}^{(\beta)}K$

$${}^{(\beta)}F^j_{ab} = R^j_{ab}+2\beta D_{[a}K^j_{b]}+\beta^2\epsilon_{jkl}K^j_a K^k_b \tag{I.1.3. 31}$$

Contracting with ${}^{(\beta)}E$ yields

$${}^{(\beta)}F^j_{ab}\ {}^{(\beta)}E^b_j = \frac{R^j_{ab}E^b_j}{\beta}+2D_{[a}(K^j_{b]}E^b_j)+\beta K^j_a G_j \tag{I.1.3. 32}$$

where we have used the Gauss constraint in the form (I.1.3. 18). We claim that the first term on the right hand side of (I.1.3. 32) vanishes identically. To see this we first derive from (I.1.3. 17) due to torsion freeness of the Christoffel connection in the language of forms the *algebraic Bianchi identity*

$$\begin{aligned}
& dx^a \wedge dx^b D_a e_b^j = de^j + \Gamma^j_k \wedge e^k = 0 \\
\Rightarrow\ & 0 = -d^2 e^j = d\Gamma^j_k \wedge e^k - \Gamma^j_l \wedge de^l = [d\Gamma^j_k + \Gamma^j_l \wedge \Gamma^l_k] \wedge e^k = \Omega^j_k \wedge e^k
\end{aligned} \tag{I.1.3. 33}$$

Now $\Omega^j_k = \Omega^i (T_i)_{jk} =: (\Omega)_{jk}$ and we see that

$$\Omega = d\Gamma + \Gamma \wedge \Gamma = d\Gamma^i\ T_i + \frac{1}{2}[T_j, T_k]\Gamma^j \wedge \Gamma^k = \frac{1}{2} dx^a \wedge dx^b R^i_{ab} T_i$$

Thus the Bianchi identity can be rewritten in the form

$$\begin{aligned}
& \epsilon_{ijk}\epsilon^{efc} R^j_{ef} e^k_c = 0 \Rightarrow \\
& \frac{1}{2}\epsilon_{ijk}\epsilon^{efc} R^j_{ef} e^k_c e^i_a = \frac{1}{2} E^b_j \epsilon_{cab}\epsilon^{efc} R^j_{ae} \\
= \ & R^j_{ab} E^b_j = 0
\end{aligned} \tag{I.1.3. 34}$$

as claimed. Now we compare with the first line of (I.1.3. 9) and thus arrive at the conclusion

$${}^{(\beta)}F^j_{ab}\ {}^{(\beta)}E^b_j = H_a + {}^{(\beta)}K^j_a G_j \tag{I.1.3. 35}$$

Next we contract (I.1.3. 36) with $\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l$ and find

$$\begin{aligned}
& {}^{(\beta)}F^j_{ab}\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l \\
= \ & \det(q)\frac{R_{abkl}e^a_k e^b_l}{\beta^2} - 2\frac{E^a_j D_a G_j}{\beta} + (K^j_a E^a_j)^2 - (K^j_b E^a_j)(K^k_a E^b_k)
\end{aligned} \tag{I.1.3. 36}$$

Expanding $v_j = e^a_j v_a,\ v_a = e^j_a v_i$, using $D_a e^j_b = 0$ and comparing $[D_a, D_b]v_j$ with $[D_a, D_b]v_c$ for any v_j we find $R_{abij} = R_{abcd}e^c_i e^d_j$ and so (I.1.3. 36) can be rewritten as

$$\begin{aligned}
& {}^{(\beta)}F^j_{ab}\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l \\
= \ & -\det(q)\frac{R}{\beta^2} - 2\ {}^{(\beta)}E^a_j D_a G_j + (K^j_a E^a_j)^2 - (K^j_b E^a_j)(K^k_a E^b_k)
\end{aligned} \tag{I.1.3. 37}$$

and comparing with the second line of (I.1.3. 9) we conclude

$$\begin{aligned}
& {}^{(\beta)}F^j_{ab}\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l + 2\ {}^{(\beta)}E^a_j D_a G_j \\
= \ & \sqrt{\det(q)}[-\sqrt{\det(q)}\frac{R}{\beta^2} - \frac{(K^j_b E^a_j)(K^k_a E^b_k) - (K^j_a E^a_j)^2}{\sqrt{\det(q)}}] \\
= \ & \frac{\sqrt{\det(q)}}{\beta^2}[-\sqrt{\det(q)}R - \beta^2\frac{(K^j_b E^a_j)(K^k_a E^b_k) - (K^j_a E^a_j)^2}{\sqrt{\det(q)}}] \\
= \ & \frac{\sqrt{\det(q)}}{\beta^2}[H + (\frac{s}{4} - \beta^2)\frac{(K^j_b E^a_j)(K^k_a E^b_k) - (K^j_a E^a_j)^2}{\sqrt{\det(q)}}] \\
= \ & 4s\sqrt{\det(q)}[-\frac{s}{4\sqrt{\det(q)}}[(K^j_b E^a_j)(K^k_a E^b_k) - (K^j_a E^a_j)^2] - \frac{s}{4\beta^2}\sqrt{\det(q)}R] \\
= \ & 4s\sqrt{\det(q)}[H - (1 + \frac{s}{4\beta^2})\sqrt{\det(q)}R]
\end{aligned} \tag{I.1.3. 38}$$

We see that the left hand side of (I.1.3. 38) is proportional to H if and only if $\beta = \pm\sqrt{s}/2$, that is, imaginary (real) for Lorentzian (Euclidean) signature. We prefer, for reasons that become obvious only in a later section, to solve (I.1.3. 38) for H as follows

$$\begin{aligned} H &= \frac{\beta^2}{\sqrt{\det({}^{(\beta)}q\beta)}}[{}^{(\beta)}F^j_{ab}\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l + 2\ {}^{(\beta)}E^a_j D_a G_j] \\ &+(\beta^2 - \frac{s}{4})\frac{({}^{(\beta)}K^j_b\ {}^{(\beta)}E^a_j)({}^{(\beta)}K^j_a\ {}^{(\beta)}E^a_j) - ({}^{(\beta)}K^j_c\ {}^{(\beta)}E^c_j)^2}{\sqrt{\det({}^{(\beta)}q\beta)}} \end{aligned} \quad \text{(I.1.3. 39)}$$

In formula (I.1.3. 39) we wrote everything in terms of ${}^{(\beta)}A, {}^{(\beta)}E$ if we understand ${}^{(\beta)}K = {}^{(\beta)}A - \Gamma$ and we used ${}^{(\beta)}q_{ab} = \beta^{-1}q_{ab} = {}^{(\beta)}E^j_a\ {}^{(\beta)}E^j_b \det({}^{(\beta)}E)$.

We notice that both (I.1.3. 35) and (I.1.3. 39) still involve the Gauss constraint. Since the transformation $K^j_a \mapsto {}^{(\beta)}A^j_a, E^a_j \mapsto {}^{(\beta)}A^j_a$ is a canonical one, the Poisson brackets among the set of first class constraints given by G_j, H_a, H are unchanged. Let us write symbolically $H_a = H'_a + f^j_a G_j, H = H' + f^j G_j$ where H'_a, H' are the pieces of H_a, H respectively not proportional to the Gauss constraint. Since G_j generates a subalgebra of the constraint algebra it follows that the modified system of constraints given by G_j, H'_a, H' not only defines the same constraint surface of the phase space but also gives a first class system again, of course, with somewhat modified algebra which however coincides with the Dirac algebra on the submanifold $G_j = 0$ of the phase space. In other words, it is completely equivalent to work with the set of constraints G_j, H'_a, H' which we write once more, dropping the prime, as

$$\begin{aligned} G_j &= {}^{(\beta)}D_a\ {}^{(\beta)}E^a_j = \partial_a\ {}^{(\beta)}E^a_j + \epsilon_{jkl}\ {}^{(\beta)}A^j_a\ {}^{(\beta)}E^a_j \\ H_a &= {}^{(\beta)}F^j_{ab}\ {}^{(\beta)}E^b_j \\ H &= [\beta^2\ {}^{(\beta)}F^j_{ab} + (\beta^2 - \frac{s}{4})\epsilon_{jmn}\ {}^{(\beta)}K^m_a\ {}^{(\beta)}K^n_b]\frac{\epsilon_{jkl}\ {}^{(\beta)}E^a_k\ {}^{(\beta)}E^b_l}{\sqrt{\det({}^{(\beta)}q\beta)}} \end{aligned} \quad \text{(I.1.3. 40)}$$

For easier comparison with the literature we also write (I.1.3. 40) in terms of ${}^{(\beta)}A^j_a, K^j_a, E^a_j$ which gives

$$\begin{aligned} G_j &= ({}^{(\beta)}D_a E^a_j)/\beta = (\partial_a\ (\beta)E^a_j + \epsilon_{jkl}\ (\beta)A^j_a E^a_j)/\beta \\ H_a &= ({}^{(\beta)}F^j_{ab}E^b_j)/\beta \\ H &= [{}^{(\beta)}F^j_{ab} + (\beta^2 - \frac{s}{4})\epsilon_{jmn}K^m_a K^n_b]\frac{\epsilon_{jkl}E^a_k E^b_l}{\sqrt{\det(q)}} \end{aligned} \quad \text{(I.1.3. 41)}$$

At this point we should say that our conventions differ slightly from those in the literature : There one writes the constraint in terms of $\tilde{K}^j_a := K^j_a/2$ and one defines ${}^{(\beta)}\tilde{K} := \beta\tilde{K} = {}^{(\beta)}K/2 = {}^{(\beta/2)}K$ and ${}^{(\beta)}\tilde{A} := \Gamma + \beta\tilde{K} = \Gamma + \beta/2K = {}^{(\beta/2)}A$ at the price of $2\ {}^{(\beta)}E$ being conjugate to ${}^{(\beta)}\tilde{A}$ instead of ${}^{(\beta)}E$ being conjugate to ${}^{(\beta)}A$. Thus ${}^{(\beta)}A = {}^{(2\beta)}\tilde{A} = {}^{(\tilde{\beta})}\tilde{A}$ with $\tilde{\beta} = 2\beta$. When writing H in terms of these quantities we find

$$H = [{}^{(\tilde{\beta})}F^j_{ab} + (\tilde{\beta}^2 - s)\epsilon_{jmn}\tilde{K}^m_a\tilde{K}^n_b]\frac{\epsilon_{jkl}E^a_k E^b_l}{\sqrt{\det(q)}} \quad \text{(I.1.3. 42)}$$

where now $\tilde{\beta}^2 = s$ is the preferred value.
Summarizing, we have rewritten the Einstein Hilbert action in the following equivalent form

$$S = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\sigma d^3x({}^{(\beta)}\dot{A}^i_a\ {}^{(\beta)}E^a_i - [\Lambda^i G_i + N^a V_a + NH]) \quad \text{(I.1.3. 43)}$$

where the appearing constraints are the ones given by either of (I.1.3. 42), (I.1.3. 41) or (I.1.3. 40).

Several remarks are in order :

- *Four-dimensional Interpretation*
Let us try to give a four-dimensional meaning to ${}^{(\beta)}A$. To that end we must complete the 3-bein e_i^a to a 4-bein e_α^μ where μ is a spacetime tensor index and $\alpha = 0,1,2,3$ an index for the defining representation of the Lorentz (Euclidean) group for $s = -1(+1)$. By definition $g_{\mu\nu}e^\mu_\alpha e^\nu_\beta = \eta_{\alpha\beta}$ is the flat Minkowski (Euclidean) metric. Thus e_0^μ, e_i^μ are orthogonal vectors and we thus choose $e_0^\mu = n^\mu$ and in the ADM frame with $\mu = t, a$ we choose $(e_i^\mu)_{\mu=a} = e_i^a$. Using the defining properties of a tetrad basis and the explicit form of $n^\mu, g_{\mu\nu}$ in the ADM frame derived earlier, above choices are sufficient to fix the tetrad components completely to be $e_0^t = 1/N,\ e_0^a = -N^a/N,\ e_i^t = 0,\ e_i^a$. Inversion gives (notice that $e_\mu^0 = se_{\mu 0} = sg_{\mu\nu}e_0^\nu = sg_{\mu\nu}n^\mu = sn_\mu$) $e_t^0 = N,\ e_a^0 = 0,\ e_t^i = N^a e_a^i, e_a^i$. Finally we have for $q_\nu^\mu = \delta_\nu^\mu - sn^\mu n_\nu = \delta_\nu^\mu - e_0^\mu e_\nu^0$ in the ADM frame $q_t^t = 0,\ q_a^t = 0,\ q_t^a = N^a, q_b^a = \delta_b^a$. Thus we obtain, modulo $G_j = 0$

$$\begin{aligned} K_a^j &= -2se_j^b K_{ab} = -2se_j^b q_a^\mu q_b^\nu \nabla_\mu n_\nu = -2e_j^b(\nabla_a e_b)^0 = 2e_j^b(\omega_a)^0{}_\alpha e_b^\alpha \\ &= 2e_j^b(\omega_a)^0{}_k e_b^k = 2(\omega_a)^0{}_j \end{aligned} \quad \text{(I.1.3. 44)}$$

where in the second identity the bracket denotes that ∇ only acts on the tensorial index and in the third we used the definition of the four dimensional spin connection $\nabla_\mu e_n u^\alpha = (\nabla_\mu e_\nu)^\alpha + (\omega_\mu)^\alpha_\beta e_n u^\beta = 0$. On the other hand we have

$$(\Gamma_a)^j{}_k e_b^k = -(D_a e_b)^j = -q_a^\mu q_b^\nu(\nabla_\mu e_\nu)^j = -(\nabla_a e_b)^j = (\omega_a)^j{}_k e_b^k \quad \text{(I.1.3. 45)}$$

whence $\omega_{ajk} = \Gamma_{ajk}$. It follows that

$${}^{(\beta)}A_{ajk} = \omega_{ajk} - 2\beta s\omega_{a0l}\epsilon_{jkl} \quad \text{(I.1.3. 46)}$$

The Hodge dual of an antisymmetric tensor $T_{\alpha\beta}$ is defined by $*T_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\beta\gamma\delta}\eta^{\gamma\gamma'}\eta^{\delta\delta'}T_{\gamma'\delta'}$. Since $\epsilon_{0ijk} = \epsilon_{ijk}$ we can write (I.1.3. 46) in the form

$${}^{(\beta)}A_{ajk} = \omega_{ajk} - 2\beta * \omega_{ajk} \quad \text{(I.1.3. 47)}$$

Now an antisymmetric tensor is called (anti)self-dual provided that $*T_{\alpha\beta} = \pm\sqrt{s}T$ with $\sqrt{s} := i^{[1-s]/2}$ and the (anti)self-dual piece of any $T_{\alpha\beta}$ is defined by $T^\pm = \frac{1}{2}[T \pm *T/\sqrt{s}]$ since $* \circ * = s$ id. An (anti)self-dual tensor therefore has only three linearly independent components. This case happens for (I.1.3. 47) provided that either $s = 1, \beta = \mp 1/2$ or $s = -1, \beta = \pm i/2$ and *in this case the new connection is just (twice) the (anti)self-dual piece of the pull-back to σ of the four-dimensional spin-connection.* In all other cases (I.1.3. 47) is only half of the information needed in order to build a four-dimensional connection and therefore we do not know how it transforms under internal boosts. This is, from this perspective, the reason why one has to gauge fix the boost symmetry of the action (I.1.3. 1) (by the time gauge $e_\mu^\alpha n^\mu = \delta_0^\alpha$) in order to remove the then present second class constraints and to arrive at the present formulation. Obviously, this is no obstacle, first, since there *does exist* a four-dimensional interpretation even in that case as we just showed and more explicitly from (I.1.3. 10) and, secondly, since we are not interested in the transformation properties under spacetime diffeomorphisms and internal Lorentz transformations of non-gauge-invariant objects anyway, although from an aesthetic point of view it would be desirable to have such an interpretation.

- *Reality Conditions*
  When β is real valued ${}^{(\beta)}A$, ${}^{(\beta)}E$ are both real valued and can directly be interpreted as the canonical pair for the phase space of an $SU(2)$ Yang-Mills theory. If β is complex then these variables are complex valued. However, they cannot be arbitrary complex functions on σ but are subject to the following reality condtions

$$ {}^{(\beta)}E/\beta = \overline{{}^{(\beta)}E/\beta},\ [{}^{(\beta)}A - \Gamma]/\beta = \overline{[{}^{(\beta)}A - \Gamma]/\beta} \tag{I.1.3. 48} $$

  where $\Gamma = \Gamma({}^{(\beta)})$ is a non-polynomial, not even analytic function. These reality conditions guarantee that there is no doubling of the number of degrees of freedom and one can check explicitly that they are preserved by the Hamiltonian flow of the constraints provided that Λ^j, the Lagrange multiplier of the Gauss constraint, is real valued. Thus, only $SU(2)$ gauge transformations are allowed but not general $SL(2,\mathbb{C})$ transformations. The reality conditions are difficult to implement in the quantum theory directly as already mentioned above.

- *Simplification of the Hamiltonian Constraint*
  The original motivation to introduce the new variables was that for the quantization of general relativity it seemed mandatory to simplify the algebraic sructure of the Hamiltonian constraint which for $s = -1$ requires $\beta = \pm i/2$ since then the constraint becomes polynomial after multiplying by a factor proportional to $\sqrt{\det(q)}$. On the other hand, then the reality conditions become non-polynomial. Finally, if one wants polynomial reality conditions then one must have β real and then the Hamiltonian constraint is still complicated. Thus it becomes questionable what has been gained. The answer is the following : For any choice of β one can actually make both the Hamiltonian constraint *and* the reality conditions polynomial by multiplying by a sufficiently high power of $\det(q)$. But the real question is whether the associated classical functions will become well-defined operator-valued distributions in quantum theory while keeping background independence. As we will see in later sections, the Hilbert space that we will choose does not support any quantum versions of these functions rescaled by powers of $\det(q)$ and there are abstract arguments that suggest that this is a representation independent statement. The requirement seems to be that the Hamiltonian constraint is a scalar density of weight one and thus we must keep the factor of $1/\sqrt{\det(q)}$ in (I.1.3. 41) whatever the choice of β and therefore the motivation for polynomiality is lost completely. The motivation to have a connection formulation rather than a metric formulation is then that *that one can go much farther in the background independent quantization programme provided that β is real.* For instance, a connection formualtion enables us to employ the powerful arsenal of techniques that have been developed for the canonical quantization of Yang-Mills theories, specifically Wilson loop techniques.

- *Choice of Fibre Bundle*
  In the whole exposition so far we have assumed that we have a trivial principal $SU(2)$ bundle over σ (see e.g. [88] for a good textbook on fibre bundle theory and section III.2) so that we can work with a globally defined connection potential and globally defined electric field ${}^{(\beta)}A$, ${}^{(\beta)}E$ respectively. What about different bundle choices ?

  Following the notation of appanedix III.2 our situation is that we are dealing with a principal $SU(2)$ bundle over σ with pull-backs ${}^{(\beta)}A_I$ by local sections of a connection and local sections ${}^{(\beta)}E_I$ of an associated (under the adjoint representation) vector bundle of two forms and would like to know whether these bundles are trivial. Since the latter is built out of the 3-beins we can equivalently look also at the frame bundle of orthonormal frames in order to decide for triviality.

Triviality of the frame bundle is equivalent to to the triviality of its associated principal bundle and in turn to σ being parallelizable. But this is automatically the case for any compact, orientable three manifold provided that $G = SU(2)$, see [89] paragraph 12, exercise 12-B. More generally, in order to prove that a principal fibre bundle is trivial one has to show that the cocycle h_{IJ} of transition functions between charts of an atlas of σ is a coboundary, that is, its (non-Abelian) Čech cohomolgy class is trivial. In [89] one uses a different method, obstruction theory, where triviality can be reduced to the vanishing of the coefficients (taking values in the homotopy groups of G) of certain cohomology groups of σ related to Stiefel-Whitney classes.

So far we did not make the assumption that σ is compact or orientable. If σ is not compact but orientable then one usually requires that there is a compact subset B of σ such that $\sigma - B$ has the topology of the complement of a ball in $\mathbb{R}^3$. Then the result holds in B and trivially in $\sigma - B$ and thus all over σ. Thus, compactness is not essential. If σ is not orientable then a smooth nowhere singular frame cannot exist and the above quaoted result does not hold, there are no smooth 3-bein fields in this case. In that case we allow non-smooth 3-bein fields, that is, we allow that $\det(e)$ has finite jumps between $\pm|\det(e)|$ on subsets of σ of Lebesgue measure zero (two surfaces) due to change of sign of one of the three forms e^j. This requires that one works with a fixed trivialization at the gauge variant level classically. At the gauge invariant level the dependence on that trivialization disappears, so there is no problem. More specifically, the constraints H, H_a as well as the symplectic structure are gauge invariant while G_j is gauge covariant so that we have independence of the choice of trivialization again on the constraint surface G_j as expected, we get equivalence with the ADM formulation.. As we will see, the choice of the bundle will become completely irrelevant anyway in the quantum theory.

- *Orientation*
  So far we did not need to impose any restriction on the orientation of the e_a^j. However, from $E^a_j = e^a_j \det(e)$ we easily obtain in $D = 3$ that $\det(E) = [\det(e)]^2 = \det(q) > 0$. Thus, classically the E^a_j are not arbitrary Lie algebra valued vector densities but rather are subject to the *anholonomic* constraint
  $$\det(E) > 0 \qquad \text{(I.1.3. 49)}$$
  One can remove this constraint by multiplying the basic variables by $\mathrm{sgn}(\det(e))$: $E^a_j := \sqrt{\det(q)}e^a_j, K^j_a = -2sK_{ab}e^b_j$ (modulo $G_j = 0$) so that in fact $\det(E) = \det(q)\mathrm{sgn}(\det(e))$ but then the result (I.1.3. 28) fails to hold (the symplectic structure remains, surprisingly, unchanged), one would get instead
  $$\int d^3x E^a_j \delta\Gamma^j_a = -\frac{1}{2}\int \mathrm{sgn}(\det(e))\epsilon^{abc}\partial_a(\delta e^j_b e^j_c) = \frac{1}{4}\int d^3x \partial_a[\mathrm{sgn}(\det(e))]\epsilon^{abc}\delta q_{bc}$$
  which is ill-defined since $0 = \epsilon^{abc}\delta q_{bc}$ is multiplied by the distribution $\partial_a[\mathrm{sgn}(\det(e))]$ unless one makes further assumptions classically such as that this distributional one form has support on a set of measure zero (motivated by the fact that q_{ab} is smooth.

In view of these considerations we will from now on only consider positive β unless otherwise specified.

I.1.4 Functional Analytic Description of Classical Connection Dynamics

In this final subsection of the classical part of this review we recall some of the elements of the usual infinite dimensional symplectic geometry that underlies gauge theories. It turns out to be rather difficult to consistently restrict the space of classical fields on a given differential manifold in such a way that the classical action remains functionally differentiable, usually critically depending on the boundary conditions that one imposes, while keeping "enough" solutions of the field equations. Usually the simplest solutions, those with a high degree of symmetry, are at the verge of lying outside of the space of fields that the variational principle was based on. Fortunately, these issues will be not too important for us as the space of quantum fields tends to be even much larger and generically is of a distributional kind without leading to any problems. Those issues will however be of some interest again when we discuss the calssical limit. We can therefore be brief here and will just sketch some of the main ideas. The interested reader is referred to the exhaustive treatment in [91].

Let G be a compact gauge group, σ a $D-$dimensional manifold which admits a principal $G-$bundle with connection over σ. Let us denote the pull-back to σ of the connection by local sections by A_a^i where $a,b,c,..=1,..,D$ denote tensorial indices and $i,j,k,..=1,..,\dim(G)$ denote indices for the Lie algebra of G. We will denote the set of all smooth connections by $\mathcal{A}$ and endow it with a globally defined metric topology of the Sobolev kind

$$d_\rho[A,A'] := \sqrt{-\frac{1}{N}\int_\sigma d^Dx\sqrt{\det(\rho)(x)}\mathrm{tr}([A_a-A'_a](x)[A_b-A'_b](x))\rho^{ab}(x)} \qquad \text{(I.1.4. 1)}$$

where $\mathrm{tr}(\tau_i\tau_j)=-N\delta_{ij}$ is our choice of normalization for the generators of a Lie algebra $Lie(G)$ of rank N and our conventions are such that $[\tau_i,\tau_j]=2f_{ij}{}^k\tau_k$ define the structure constants of $Lie(G)$. Here ρ_{ab} is a fiducial metric on σ of everywhere Euclidean signature. In what follows we assume that either $D\neq 2$ (for $D=2$, (I.1.4. 1) depends only on the conformal structure of ρ and cannot guarantee convergence for arbitrary fall-off conditions on the connections) or that $D=2$ and the fields A are Lebesgue integrable.

Let F_j^a be a Lie algebra valued vector density test field of weight one and let f_a^j be a Lie algebra valued covector test field. Let, as before A_a^j be a the pull-back of a connection to σ and consider a vector bundle of electric fields, that is, of Lie algebra valued vector densities of weight one whose bundle projection to σ we denote by E_i^a. We consider the smeared quantities

$$F(A):=\int_\sigma d^Dx F_i^a A_a^i \text{ and } E(f):=\int_\sigma d^Dx E_i^a f_a^i \qquad \text{(I.1.4. 2)}$$

While both are diffeomorphism covariant it is only the latter which is gauge covariant, one reason to consider the singular smearing through holonomies discussed below. The choice of the space of pairs of test fields $(F,f)\in\mathcal{S}$ depends on the boundary conditions on the space of connections and electric fields which in turn depends on the topology of σ and will not be specified in what follows.

We now want to select a subset $\mathcal{M}$ of the set of all pairs of smooth functions (A,E) on σ such that (I.1.4. 2) is well defined (finite) for any $(F,f)\in\mathcal{S}$ and endow it with a manifold structure and a symplectic structure, that is, we wish to turn it into an infinite dimensional symplectic manifold.

We define a topology on $\mathcal{M}$ through the metric:

$$d_{\rho,\sigma}[(A,E),(A',E')] \qquad \text{(I.1.4. 3)}$$
$$:= \sqrt{-\frac{1}{N}\int_\sigma d^Dx[\sqrt{\det(\rho)}\rho^{ab}\mathrm{tr}([A_a-A'_a][A_b-A'_b])+\frac{\sigma_{ab}\mathrm{tr}([E^a-E^{a\prime}][E^b-E^{b\prime}])}{\sqrt{\det(\sigma)}}]}$$

where ρ_{ab},σ_{ab} are again fiducial metrics on σ of everywhere Euclidean signature. Their fall-off behaviour has to be suited to the boundary conditions of the fields A,E at spatial infinity. Notice that the metric (I.1.4. 3) is gauge invariant (and thus globally defined, i.e. is independent of the choice of local section) and diffeomorphism covariant and that $d_{\rho,\sigma}[(A,E),(A',E')] = d_\rho[A,A'] + d_\sigma[E,E']$ (recall (I.1.1. 1)).

Now, while the space of electric fields in Yang-Mills theory is a vector space, the space of connections is only an affine space. However, as we have also applications in general relativity with asymptotically Minkowskian boundary conditions in mind, also the space of electric fields will in general not be a vector space. Thus, in order to induce a norm from (I.1.4. 3) we proceed as follows: Consider an atlas of $\mathcal{M}$ consisting only of M itself and choose a fiducial background connection and electric field $A^{(0)}, E^{(0)}$ (for instance $A^{(0)} = 0$). We define the global chart

$$\varphi\;:\;\mathcal{M} \mapsto \mathcal{E};\;(A,E) \mapsto (A - A^{(0)}, E - E^{(0)}) \qquad \text{(I.1.4. 4)}$$

of $\mathcal{M}$ onto the vector space of pairs $(A - A^{(0)}, E - E^{(0)})$. Obviously, φ is a bijection. We topologize $\mathcal{E}$ in the norm

$$||(A - A^{(0)}, E - E^{(0)})||_{\rho\sigma} := \sqrt{d_{\rho\sigma}[(A,E),(A^{(0)},E^{(0)})]} \qquad \text{(I.1.4. 5)}$$

The norm (I.1.4. 5) is of course no longer gauge and diffeomorphism covariant since the fields $A^{(0)}, E^{(0)}$ do not transform, they are background fields. We need it, however, only in order to encode the fall-off behaviour of the fields which are independent of gauge – and diffeomorphism covariance.

Notice that the metric induced by this norm coincides with (I.1.4. 3). In the terminology of weighted Sobolev spaces the completion of $\mathcal{E}$ in the norm (I.1.4. 5) is called the Sobolev space $H^2_{0,\rho} \times H^2_{0,\sigma^{-1}}$, see e.g. [92]. We will call the completed space $\mathcal{E}$ again and its image under φ^{-1}, $\mathcal{M}$ again (the dependence of φ on $(A^{(0)}, E^{(0)})$ will be suppressed). Thus, $\mathcal{E}$ is a normed, complete vector space, that is, a Banach space, in fact it is even a Hilbert space. Moreover, we have modelled $\mathcal{M}$ on the Banach space $\mathcal{E}$, that is, $\mathcal{M}$ acquires the structure of a (so far only topological) Banach manifold. However, since $\mathcal{M}$ can be covered by a single chart and the identity map on $\mathcal{E}$ is certainly C^∞, M is actually a smooth manifold. The advantage of modelling $\mathcal{M}$ on a Banach manifold is that one can take over almost all the pleasant properties from the finite dimensional case to the infinite dimensional one (in particular, the inverse function theorem).

Next we study differential geometry on $\mathcal{M}$ with the standard techniques of calculus on infinite dimensional manifolds (see e.g. [93]). We will not repeat all the technicalities of the definitions involved, the interested reader is referred to the literature quoted.

i) A function $f\,:\;\mathcal{M} \mapsto \mathbb{C}$ on $\mathcal{M}$ is said to be differentiable at m if $g := f \circ \varphi^{-1}\,:\;\mathcal{E} \mapsto \mathbb{C}$ is differentiable at $u = \varphi(m)$, that is, there exist *bounded* linear operators $Dg_u, Rg_u\;:\;\mathcal{E} \mapsto \mathbb{C}$ such that

$$g(u+v) - g(u) = (Dg_u)\cdot v + (Rg_u)\cdot v \text{ where } \lim_{||v||\to 0} \frac{|(Rg_u)\cdot v|}{||v||} = 0\;. \qquad \text{(I.1.4. 6)}$$

$Df_m := Dg_u$ is called the functional derivative of f at m (notice that we identify, as usual, the tangent space of $\mathcal{M}$ at m with $\mathcal{E}$). The definition extends in an obvious way to the case where $\mathbb{C}$ is replaced by another Banach manifold. The equivalence class of functions differentiable at m is called the germ $G(m)$ at m. Here two functions are said to be equivalent provided they coincide in a neighbourhood containing m.

ii) In general, a tangent vector v_m at $m \in \mathcal{M}$ is an equivalence class of triples (U, φ, v_m) where (U,φ) is a chart of the atlas of $\mathcal{M}$ containing m and $v_m \in \mathcal{E}$. Two triples are said to be equivalent provided that $v'_m = D(\varphi' \circ \varphi^{-1})_{\varphi(m)} \cdot v_m$. In our case we have only one chart and equivalence becomes trivial. Tangent vectors at m can be considererd as derivatives on the germ $G(m)$ by defining

$$v_m(f) := (Df_m) \cdot v_m = (D(f \circ \varphi^{-1})_{\varphi(m)}) \cdot v_m \tag{I.1.4. 7}$$

Notice that the definition depends only on the equivalence class and not on the representative. The set of vectors tangent at m defines the tangent space $T_m(\mathcal{M})$ of $\mathcal{M}$ at m.

iii) The cotangent space $T'_m(\mathcal{M})$ is the topological dual of $T_m(\mathcal{M})$, that is, the set of *continuous* linear functionals on $T_m(\mathcal{M})$. It is obviously isomorphic with $\mathcal{E}'$, the topological dual of $\mathcal{E}$. Since our model space $\mathcal{E}$ is reflexive (it is a Hilbert space) we can naturally identify tangent and cotangent space (by the Riesz lemma) which also makes the definition of contravariant tensors less ambiguous. We will, however, not need them for what follows. Similarly, one defines the space of $p-$covariant tensors at $m \in \mathcal{M}$ as the space of *continuous* $p-$linear forms on the $p-$fold tensor product of $T_m(\mathcal{M})$.

iv) So far the fact that $\mathcal{E}$ is a Banach manifold was not very crucial. But while the tangent bundle $T(\mathcal{M}) = \cup_{m\in\mathcal{M}} T_m(\mathcal{M})$ carries a natural manifold structure modelled on $\mathcal{E}\times\mathcal{E}$ for a general Fréchet space (or even locally convex space) $\mathcal{E}$ the cotangent bundle $T'(M) = \cup_{m\in\mathcal{M}} T'_m(\mathcal{M})$ carries a manifold structure only when $\mathcal{E}$ is a Banach space as one needs the inverse function theorem to show that each chart is not only a differentiable bijection but that also its inverse is differentiable. In our case again there is no problem. We define differentiable vector fields and $p-$covariant tensor fields as cross sections of the corresponding fibre bundles.

v) A differential form of degree p on $\mathcal{M}$ or $p-$form is a cross section of the fibre bundle of completely skew continuous $p-$linear forms. Exterior product, pull-back, exterior differential, interior product with vector fields and Lie derivatives are defined as in the finite dimensional case.

Definition I.1.1 *Let $\mathcal{M}$ be a differentiable manifold modelled on a Banach space $\mathcal{E}$. A weak respectively strong symplectic structure Ω on M is a closed 2-form such that for all $m \in \mathcal{M}$ the map*

$$\Omega_m :\ T_m(\mathcal{M}) \mapsto T'_m(\mathcal{M});\ v_m \to \Omega(v_m, .) \tag{I.1.4. 8}$$

is an injection respectively a bijection.

Strong symplectic structures are more useful because weak symplectic structures do not allow us to define Hamiltonian vector fields through the definition $DL + i_{\chi_L}\Omega = 0$ for differentiable L on M and Poisson brackets through $\{f, g\} := \Omega(\chi_f, \chi_g)$, see e.g. [94] for details.

Thus we define finally a strong symplectic structure for our case by

$$\Omega((f, F), (f', F')) := \int_\Sigma d^D x [F^a_i f^{i\prime}_a - F^{a\prime}_i f^i_a](x) \tag{I.1.4. 9}$$

for any $(f, F), (f', F') \in \mathcal{E}$. To see that Ω is a strong symplectic structure we observe first that the integral kernel of Ω is constant so that Ω is clearly exact, so, in particular, closed. Next, let $\theta \in \mathcal{E}' \equiv \mathcal{E}$. To show that Ω is a bijection it suffices to show that it is a surjection (injectivity follows trivially from linearity). We must find $(f, F) \in \mathcal{E}$ so that $\theta(.) = \Omega((f, F), .)$ for any one-form θ . Now by the Riesz lemma there exists $(f_\theta, F_\theta) \in \mathcal{E}$ such that $\theta(.) =< (f_\theta, F_\theta), . >$ where $< ., . >$ is the inner product induced by (I.2.1. 4). Comparing (I.1.4. 3) and (I.1.4. 9) we see that we have achieved our goal provided that the functions

$$F_i^a := \rho^{ab}\sqrt{\det(\rho)} f_{b\theta}^i,\ f_a^i := -\frac{\sigma_{ab}}{\sqrt{\det(\rho)}} F_{i\theta}^b \tag{I.1.4. 10}$$

are elements of $\mathcal{E}$. Inserting the definitions we see that this will be the case provided that the functions $\rho^{cd}\sigma_{ca}\sigma_{db}/\sqrt{\det(\rho)}$ and $\det(\rho)\sigma_{cd}\rho^{ca}\rho^{db}/\sqrt{\det(\sigma)}$ respectively fall off at least as $\sigma_{ab}/\sqrt{\det(\sigma)}$ and $\rho^{ab}\sqrt{\det(\rho)}$ respectively. In physical applications these metrics are usually chosen to be of the form $1+O(1/r)$ where r is an asymptotical radius function so that these conditions are certainly satisfied. Therefore, $(f,F)\in\mathcal{E}$ and our small lemma is established.

Let us compute the Hamiltonian vector field of a function L on our $\mathcal{M}$. By definition for all $(f,F)\in\mathcal{E}$ we have at $m=(A,E)$

$$DL_m\cdot(f,F)=\int_\Sigma d^Dx[(DL_m)_i^a f_a^i+(DL_m)_a^i F_i^a]=-\int_\Sigma d^Dx[(\chi_{Lm})_i^a f_a^i-(\chi_{Lm})_a^i F_i^a] \tag{I.1.4. 11}$$

thus $(\chi_L)_i^a=-(DL)_i^a$ and $(\chi_L)_a^i=(DL)_a^i$. Obviously, this defines a bounded operator on $\mathcal{E}$ if and only if L is differentiable. Finally, the Poisson bracket is given by

$$\{L,L'\}_m=\Omega(\chi_L,\chi_{L'})=\int_\Sigma d^Dx[(DL_m)_a^i(DL'_m)_i^a-(DL_m)_i^a(DL'_m)_a^i] \tag{I.1.4. 12}$$

It is easy to see that Ω has the symplectic potential Θ, a one-form on $\mathcal{M}$, defined by

$$\Theta_m((f,F))=\int_\Sigma d^Dx E_i^a f_a^i \tag{I.1.4. 13}$$

since

$$D\Theta_m((f,F),(f',F')):=(D(\Theta_m)\cdot(f,F))\cdot(f',F')-(D(\Theta_m)\cdot(f',F'))\cdot(f,F)$$

and $DE_i^a(x)_m\cdot(f,F)=F_i^a(x)$ as follows from the definition.

Coming back to the choice of $\mathcal{S}$, it will in general be a subspace of $\mathcal{E}$ so that (I.2.0. 15) still converges. We can now compute the Poisson brackets between the functions $F(A), E(f)$ on M and find

$$\{E(f),E(f')\}=\{F(A),F'(A)\}=0,\ \{E(f),A(F)\}=F(f) \tag{I.1.4. 14}$$

Remark :
In physicists' notation one often writes $(DL_m)_a^i(x):=\frac{\delta L}{\delta A_a^i(x)}$ etc. and one writes the symplectic structure as $\Omega=\int d^Dx\ DE_i^a(x)\wedge DA_a^i(x)$.

I.2 Mathematical Foundations of Modern Canonical Quantum General Relativity

In the previous section we have derived a canonical connection formulation of classical general relativity. We have emphasized the importance of n-form fields for a background independent quantization of the theory. In this section we will see that the insistence on background independent methods results in a Hilbert space that is drastically different from the usual Fock space employed in perturbative quantum field theory.

We begin by sketching the history of the subject:
In the previous section we have shown that for $\beta = \pm i, s = -1$ the Hamiltonian constraint greatly simplifies, up to a factor of $1/\sqrt{\det(q)}$ it becomes a fourth order polynomial in ${}^{\mathbb{C}}A_a^j, E^a_j$. In order to find solutions to the quantum constraint one chose a *holomorphic* connection representation, that is, wave functions are functionals of ${}^{\mathbb{C}}A$ but not of $\overline{{}^{\mathbb{C}}A}$, the connection itself becomes a multiplication operator while the electric field becomes a functional differential operator. In formulae for the choice $\beta = -i$,

$$({}^{\mathbb{C}}\hat{A}_a^j(x)\psi)[{}^{\mathbb{C}}A] = {}^{\mathbb{C}}A_a^j(x)\psi[{}^{\mathbb{C}}A] \text{ and } (\hat{E}^a_j(x)\psi)[{}^{\mathbb{C}}A] = \ell_p^2 \frac{\delta\psi[{}^{\mathbb{C}}A]}{\delta {}^{\mathbb{C}}A_a^j(x)} \quad \text{(I.2.0. 15)}$$

(notice that iE/κ is conjugate to ${}^{\mathbb{C}}A$, $\ell_p^2 = \hbar\kappa$ is the Planck area). With this definition, which is only formal at this point since one does not know what the functional derivative means without specifying the function space to which the ${}^{\mathbb{C}}A$ belong, the canonical commutation relations

$$[{}^{\mathbb{C}}\hat{A}_a^j(x), {}^{\mathbb{C}}\hat{A}_b^k(y)] = [\hat{E}^a_j(x), \hat{E}^b_k(y)] = 0, \; [\hat{E}_a^j(x), {}^{\mathbb{C}}\hat{A}_b^k(y)] = \ell_p^2 \delta_b^a \delta_j^k \delta(x,y) \quad \text{(I.2.0. 16)}$$

are formally satisfied. However, the adjointness relations

$$(\hat{E}_a^j(x))^\dagger = \hat{E}_a^j(x), \; {}^{\mathbb{C}}\hat{A}_a^j(x) + ({}^{\mathbb{C}}\hat{A}_a^j(x))^\dagger = 2\hat{\Gamma}_a^j(x) \quad \text{(I.2.0. 17)}$$

could not be checked because no scalar product was defined with respect to which (I.2.0. 17) should hold. Besides simpler mathematical problems such as domains of definitions of the operator valued distributions (I.2.0. 15), equation (I.2.0. 17) looks desastrous in view of the explicit formula (I.1.3. 20) for the spin connection where operator valued distributions would appear multiplied not only at the same point but also in the denominator which would be extremely difficult to define if possible at all and could prevent one from defining a positive definite scalar product with respect to which the adjointness conditions should hold.

The implementation of the adjointness relations (which one can make polynomial by multiplying ${}^{\mathbb{C}}A$ by a sufficiently high power of the operator corresponding to $\det(q)$) continues to be the major obstacle with the complex connection formulation even today which, is why the real connection formulation is favoured at the moment. However, in these pioneering years at the end of the 90's nobody thought about using real connections since the simplification of the Hamiltonian constraint seemed to be the most important property to preserve which is why researchers postponed the solution of the adjointness relations and the definition of an inner product to a later stage and focussed first on other problems. There was no concrete proposal at that time how to do that but the fact that the complex connection ${}^{\mathbb{C}}A = \Gamma - iK$ is reminecent of the harmonic oscillator variable $z = x - ip$ made it plausible that one could possibly make use of the technology known from geometric quantization concerning complex Kähler polarizations [51] and the relevant Bargmann-Segal transformation theory. A concrete proposal in terms of the phase space Wick rotation transformation mentioned earlier appeared only later in [80] but until today these ideas have not been mathematically rigorously implemented.

To speak with the critics of the connection variable approach[2] "The new variable people have a credit card that is called "Adjointness Relations". Whenever they meet a problem that they cannot solve, they charge the credit card. But one day they must pay the price for their charges and I wonder what will happen then".

If one would always first worry about the potential problems and not sometimes close one's eyes and push foreward anyway, then progress would never have been made in physics. There was a mutitude of results that one could obtain by formal manipulations even in absence of an inner product. The most important observation at that time, in the opinion of the author, is the discovery of the importance of the use of holonomy variables also known as Wilson loop functions. We will drop the superscripts $\beta, \mathbb{C}$ in what follows.

Already in the early 80's Gambini et al [95] pointed out the usefulness of Wilson-loop functions for the canonical quantization of Yang-Mills theory. Given a directed loop (closed path) α in σ and a G-connection A for some gauge group G one can consider the holonomy $h_\alpha(A)$ of A along α. The holonomy of a connection is abstractly defined via principal fibre bundle theory, but physicists prefer the formula $h_\alpha(A) := \mathcal{P}\exp(\oint_\alpha A)$ where $\mathcal{P}$ stands for path-ordering the power expansion of the exponential in such a way that the connection variables are ordered from left to right with the parameter along the loop on which they depend increasing. We will give a precise definition later on. The connection can be taken in any representation of G but we will be mostly concerned with $G = SU(2)$ and will choose the fundamental representation (in case of $G = SL(2, \mathbb{C})$ one chooses one of its two fundamental representations). The Wilson loop functions are then given by

$$T_\alpha(A) := \mathrm{tr}(h_\alpha(A)) \tag{I.2.0. 18}$$

where tr denotes the corresponding trace. The importance of such Wilson loop functions is that, at least for compact groups, one knows that they capture the full gauge invariant information about the connection [96]. For the case at hand, $SL(2, \mathbb{C})$, an independent proof exists [97].

After the introduction of the new variables which display general relativity as a special kind of Yang-Mills theory, Jacobson, Rovelli and Smolin independently rediscovered and applied Gambini's ideas to canonical quantum gravity [98]. Since the connection representation was holomorphic, one needed only one of the fundamental represenations of $SL(2, \mathbb{C})$ (and not its complex conjugate).

The author does not want to go very much into details about the rich amount of formal and exact results that were obtained by working with these loop variables before 1992 but just list the most important ones. An excellent review of these issues is contained in the book by Gambini and Pullin [99] which has become the standard introductory reference on the loop representation.

1) *Formal Solutions to the Hamiltonian Constraint in the Connection Representation*
   By ordering the operators $\hat{E}$ to the right in the quantization of the rescaled density weight two operator corresponding to $\tilde{H}$, one can show [98] that formally $\hat{\tilde{H}}T_\alpha = 0$ for every non-intersecting smooth loop α (see also [100] for an extension to more complicated loops). The formal character of this argument is due to the fact that this is a regulated calculation where in the limit as the regulator is removed one multiplies zero by infinity. An important role plays the notion of a so-called "area-derivative".

2) *Loop Transform and Knot Invariants*
   Since the diffeomorphism constraint maps a Wilson loop function to a Wilson loop function for a diffeomorphic loop one immediately sees that knot invariants should play an important

[2] Quotation from a comment given by Karel Kuchař to the author just before his talk at the meeting "Quantum Gravity in the Southern Cone", Punta del Este, Uruguay, 1996.

role. Let μ be a diffeomorphism invariant measure on some space of connections, α a loop and ψ any state. One can then define a loop transformed state by $\psi'(\alpha) := \int d\mu(A)\overline{T_\alpha(A)}\psi(A)$. The state $\Psi = 1$ is annihilated by the diffeomorphism constraint if we define the action of an operator $\hat{O}'$ in this loop representation by $(\hat{O}'\psi')(\alpha) := \int d\mu(A)\overline{(\hat{O}T_\alpha)(A)}\psi(A)$ where $\hat{O}$ is its action in the connection representation. Likewise one sees, at least formally, that if α is a smooth non-self-intersecting loop then $\psi'(\alpha)$ is annihilated by the Hamiltonian constraint. Of course, again this is rather formal because a suitable diffeomorphism invariant measure μ was not known to exist.

3) *Chern-Simons Theory*
If one considers, in particular, the loop transform with respect to the formal measure given by Lebesgue measure times the exponential of i/λ times the Chern-Simons action where λ is the cosmological constant then one can argue to obtain particular knot invariants related to the Jones-polynomial [99, 101] the coefficients of which seem to be formal solutions to the Hamiltonian constraint in the loop representation with a cosmological term: Since the exponential of the Chern-Simons action is also a formal solution to the Hamiltonian constraint with a cosmological term in the connection representation [102] with momenta ordered to the left one obtains solutions to the Hamiltonian constraint (provided a certain formal integration by parts formula holds) which correspond to arbitrary, possibly intersecting, loops.

4) *Commutators*
Also commutators of constraints were studied formally in the loop representation reproducing the Poisson algebra up to quantities which become singular as the regulator is removed (see [99]). These singular coefficients will later be seen to come from the fact that $\tilde{H}$ is a density of weight two rather than one. Such singularities must be removed but this could be done for $\tilde{H}$ only by breaking diffeomorphism invariance which is unacceptable in quantum gravity. We will come back to this point later.

5) *Model Systems*
One could confirm the validity of the connection representation in exactly solvable model systems such as the familiar mini – and midisuperspace models based on Killing – or dimensional reduction for which the reality conditions can be addressed and solved quantum mechanically [103].

These developments in the years 1987-92 confirmed that using Wilson loop functions was something extremely powerful and a rigorous quantization of the theory should be based on them. Unfortunately, all the nice results obtained so far in the full theory, especially concerning the dynamics as, e.g., the existence of solutions to the constraints, were only formal because there was no Hilbert space available which would enable one to say in which topology certain limits might exist or not.

The time had come to invoke rigorous functional analysis into the approach. Unfortunately, this was not possible so far for quantum theories of connections for non-compact gauge groups such as $SL(2,\mathbb{C})$ but only for arbitrary compact gauge groups. The motivation behind pushing these developments anyway at that time had been, again, that by using Bargmann-Segal transformation theory one would be able to transfer the results obtained to the physically interesting case. Luckily, due to the results of [78] one could avoid this aditional step and make the results of this section directly available for Lorentzian quantum gravity, although in the real connection formulation rather than the complex one.

i) *1992: Quantum Configuration Space*
The first functional analytic ideas appeared in the seminal paper by Ashtekar and Isham [104] in which they constructed a quantum configuration space of distributional connections $\overline{\mathcal{A}}$ by using abstract Gel'fand – Naimark – Segal (GNS) theory for Abelean C^* algebras, see appendix III.4. In quantum field theory it is generic that the measure underlying the scalar product of the theory is supported on a distributional extension of the classical configuration space and therefore it was natural to look for something similar, although in a background independent context. Rendall [105] was able to show that the classical configuration space of smooth connections $\mathcal{A}$ is topologically densely embedded into $\overline{\mathcal{A}}$.

ii) *1993 – 1994: Measure Theory, Projective Techniques*
Ashtekar and Lewandowski [106] then succeeded in providing $\overline{\mathcal{A}}$ with a σ-algebra of measurable subsets of $\overline{\mathcal{A}}$ and giving a cylindrical defintion of a measure μ_0 which is invariant under G gauge transformations and invariant under the spatial diffeomorphisms of Diff(σ). In [107] Marolf and Mourão established that this cylindrically defined measure has a unique σ-additive extension to the just mentioned $\sigma-$algebra. Moreover, they proved that, expectedly, $\mathcal{A}$ is contained in a measurable subset of $\overline{\mathcal{A}}$ of measure zero and introduced projective techniques into the framework. In [108] Ashtekar and Lewandowski developed the projective techniques further and used them in [109] to set up integral and differential calculus of on $\overline{\mathcal{A}}$.

Also Baez [110] had constructed different spatially diffeomorphism invariant measures on $\overline{\mathcal{A}}$, however, they are not faithful (do not induce positive definite scalar products).

iii) *1994: Complex Connections and Heat Kernel Measures*
The Segal-Bargmann representation in ordinary quantum mechanics on the phase space $\mathbb{R}^2$ is a representation in which wave functions are holomorphic, square integrable (with respect to the Liouville measure) functions of the complex variable $z = q - ip \in \mathbb{C}$. One can obtain this representation by heat kernel evolution followed by analytic continuation from the usual position space representation. In [79] Hall generalized this unitary, so-called Segal-Bargmann transformation, to phase spaces which are cotangent bundles over arbitrary compact gauge groups based on the observation that a natural Laplace operator (generator of the heat kernel evolution) exists on such groups. The role of $\mathbb{C}$ is then replaced by the complexification $G^{\mathbb{C}}$ of G. Since it turns out that the Hilbert space of functions on $\overline{\mathcal{A}}$ labelled by a piecewise analytical loop reduces to $SU(2)^N$ for some finite natural number N one can just apply Hall's construction to quantum gravity which would seem to map us from the real connection representation to the complex one. This was done in [83] by Ashtekar, Lewandowski, Marolf, Mourão and Thiemann. The question remained wether the so obtained inner product incorporates the correct adjointness – and canonical commutation relations among the complexified holonomies. In [80, 81] this was shown not to be the case but at the same time a proposal was made for how to modify the transform in such a way that the correct adjointness – and canonical commutation relations are guaranteed to hold. This so-called Wick rotation transformation is a special case of an even more general method, the so-called complexifier method, which consists in replacing the Laplacian by a more general operator (the complexifier) and can be utilized, as in the case of quantum gravity, to keep the algebraic structure of an operator simple while at the same time trivializing the adjointness conditions on the inner product. Unfortunately, the Wick rotation generator for quantum gravity is very complicated which is why there is no rigorous proof to date for the existence and the unitarity of the proposed transform.

iv) *1995: Hilbert Space, Adjointness Relations and Canonical Commutation Relations*

In [54] Ashtekar, Lewandowski, Marolf, Mourão and Thiemann could show that the Hilbert space $\mathcal{H}_0 = L_2(\overline{\mathcal{A}}, d\mu_0)$ in fact solves the adjointness – and canonical commutation relations for any canonical quantum field theory of connections that is based on a comapct gauge group provided one represents the connection as a multiplication operator and the electric field as a functional derivative operator, in fact, with these actions of the operators the measure μ_0 on $\overline{\mathcal{A}}$ is almost uniquely selected. The results of [54] demonstrated that the Hilbert space $\mathcal{H}_0$ provides in fact a physically correct, kinematical representation for such theories. Kinematical here means that the elements of this Hilbert space carry a representation of the constraint operators but are not annihilated by them, that is, they are not physical (or dynamical) states. These authors were also able to provide, in the same paper, the complete set of solutions of the spatial, analytic diffeomorphism constraint (labelled by singular (intersecting) knot classes) plus a physical (with respect to the spatial diffeomorphism constraint) inner product using group averaging methods [56] (Gel'fand triple techniques) and thus, as a side result, showed that the Husain-Kuchař [111] model is a completely integrable, diffeomorphism invariant quantum field theory. Group averaging methods provide a sytematic framework of how to go from the kinematical Hilbert space to the physical one.

v) *1995: Loop – and Connection Representation: Spin Network Functions*
Quite independently, Rovelli and Smolin as well as as Gambini and Pullin et al had pushed another representation of the canonical commutation relations, the so-called loop representation already mentioned above for which states of the Hilbert space are to be thought of as functionals of loops rather than connections. Since the Wilson loop functionals (polynomials of traces of holonomies) are not linearly independent, they are subject to the so-called Mandelstam identities, it was mandatory to first find a set of linearly independent functions. Using older ideas due to Penrose [112] Rovelli and Smolin [113] were able to write down such loop functionals, later called spin-network functions, that are labelled by a smooth $SU(2)$ connection. They then introduced an inner product between these functions by simply *defining* them to be orthonormal. Baez [114] then proved that, using that spin-network functions (considered as functionals of connections labelled by loops) can in fact be extended to $\overline{\mathcal{A}}$, the spin-network functions are indeed orthonormal with respect to $\mathcal{H}_0$, moreover, they form a basis, the two Hilbert spaces defined by Ashtekar and Lewandowski on the one hand and Rovelli and Smolin are indeed unitarily equivalent. In [115] Thiemann proved a Plancherel theorem, saying that, expectedly, the loop representation and the connection representation are like mutual, non-Abelean Fourier transforms (called the loop transform as mentioned above) of each other where the role of the kernel of the transform is played by the spin-network functions as one would intuitively expect because they are labelled by both loops and connections.

vi) *1996 – 1998: Analytical Versus Smooth and Piecewise Linear Loops*
In all these developments it was crucial, for reasons that will be explained below, that σ is an analytic manifold and that the loops were piecewise analytic. Baez and Sawin [116] were able to transfer much of the structure to the case that the loops are only piecewise smooth and intersect in a controlled way (a so-called web) and some of their results were strengthened by Lewandowski and Thiemann [117]. In [118] Zapata introduced the concept of piecewise linear loops. The motivation for these modifications was that the analytical category is rather unnatural from a physical viewpoint although it is a great technical simplification. For instance, in the smooth category there is no spin network basis any longer. Both in the analytic and smooth category the Hilbert space is non-separable after moding out by analytic or smooth diffeomorphisms respectively while in the piecewise linear category one ends up with a separable

Hilbert space. The motivation for the piecewise linear category is, however, unclear from a classical viewpoint (for instance the classical action is not invariant under piecewise linear diffeomorphisms). In [119] arguments were given, that support the fact that the (mutually orthogonal, unitarily equivalent) Hilbert spaces labelled by the continuous moduli that still appear in the diffeomorphism invariant analytic and smooth category are superselected. If one fixes the moduli, the Hilbert space becomes separable.

vii) *1994 – 2001: Relation with Constructive Quantum (Gauge) Field Theory*
One may wonder whether the techniques associated with $\overline{\mathcal{A}}$ can be applied to ordinary Yang-Mills theory on a background metric. The rigorous quantization of Yang-Mills theory on Minkowski space is still one of the major challenges of theoretical and mathematical physics [120]. There is a vast literature on this subject [121] and the most advanced results in this respect are undobtedly due to Balaban et al which are so difficult to understand "... that they lie beyond the limits of human communicational abilities..." [122]. Technically the problem has been formulated in the context of constructive (Euclidean) quantum field theory [32] which is geared to scalar fields propagating on Minkowski space. In [123] a proposal for a generalization of the key axioms of the framework, the so-called Osterwalder-Schrader axioms [124], has been given by Ashtekar, Lewandowski, Marolf, Mourão and Thiemann. These were then successfully applied in [125] by the same authors to the completely solvable Yang-Mills theory in two dimensions by making explicit use of $\overline{\mathcal{A}}, \mu_0$ and spin-network techniques which so far had not been done before although the literature on Yang-Mills theory in two dimensions is rather vast [126]. These results have been refined by Fleischhack [127]. It became clear these axioms apply only to background independent gauge field theories which is why it works in two dimensions only (in two dimensions Yang Mills theory is not background independent but almost: it is invariant under area preserving diffeomorphisms which turns out to be sufficient for the constructions to work out). This motivated Ashtekar, Marolf, Mourão and Thiemann to generalize the Osterwalder-Schrader framework to general diffeomorphism invariant quantum field theories [128]. Surprisingly the key theorem of the whole approach, the Osterwalder-Schrader reconstruction theorem that allows to obtain the Hilbert space of the canonical quantum field theory from the Euclidean one, can be straightforwardly adapted to the more general context.

One of the Osterwalder-Schrader axioms is the uniqueness of the vacuum which is stated in terms of the ergodicity property of the underlying measure with respect to the time translation subgroup of the Euclidean group (see e.g. [129]) which in turn has consequences for the support properties of the measure. In [130] Mourão, Velhinho and Thiemann analyzed these issues for μ_0 and found ergodicity with respect to any infinite, discrete subgroup of the diffeomorphism group which implied a refinement of the support properties established in [107].

viii) *1999 – 2001: Categories and Groupoids, Hyphs and Gauge Orbit Structure of* $\overline{\mathcal{A}}$
Following an earlier idea due to Baez [131] Velhinho [132] gave a nice categorical and purely algebraic chracterization of $\overline{\mathcal{A}}$ and all the structure that comes with it without using C^* techniques. The technical simplifications that are involved rest on the concept of a groupoid of piecewise analytic paths in σ rather than (base-pointed) loops.

In [133] Fleischhack, motivated by his results in [127], discussed a new notion of "loop independence" which has the advantage of being independent of the differentiability category of the graphs under consideration and in particular includes the analytical and smooth category. The new type of collections of loops are called hyphs. A hyph is a finite collection of piecewise C^r paths together with an ordering $\alpha \mapsto p_\alpha$ of its paths p_α where α belongs to some linearly ordered index set such that p_α is independent of all the paths $\{p_\beta; \beta < \alpha\}$. Here a path p is said to be independent of another path p' if there exists a free point x on p (which may be one of its boundary points), that is, there is a segment of p incident at x which does not overlap with a segment of p' (although p, p' may intersect in x). That is, path independence is based on the germ of a path. In contrast to graphs or webs (collections of piecewise analytical or smooth paths), a hyph requires an ordering. Nevertheless one can get as far with hyphs as with webs but not as far as with graphs.

Fleischhack also investigated the issue of Gribov copies in $\overline{\mathcal{A}}$ [134] with respect to $SU(2)$ gauge transformations. It should be noted that fortunately Gribov copies are no problem in our context: The measure is a probability measure and the gauge group therfore has finite volume. Integrals over gauge invariant functions are therefore well-defined.

ix) *2000 – 2001: Infinite Tensor Product Extension*
Finally, the Hilbert space $\mathcal{H}_0$ is sufficient for the applications of quantum general relativity only if σ is compact. In the non-compact case an extension from compactly supported to non-compactly supported, piecewise analytic paths becomes necessary. Thiemann and Winkler [136] discovered that the framework of the Infinite Tensor Product of Hilbert spaces, developed by von Neumann more than 60 years ago, is ideally suited to deal with this problem. In contrast to $\mathcal{H}_0$ the extended Hilbert space $\mathcal{H}^\otimes$ is no L_2 space any longer.

We notice that all these developments still use a concrete manifold σ and that the loops or paths are embedded into it. However, in order to describe topology change within quantum gravity it would be desirable to formulate a Hilbert space using non-embedded (algebraic) graphs [137]. The state of the abstract Hilbert space itself should tell us into wich σ's the algebraic graph on which it is based can be embedded. For some ideas into that direction in connection with semiclassical issues see [138].

This concludes our historical overview over the development of the subject. In the next section we try to give a modern introduction into the key structural theorems by combining most of the above cited literature which means that we will depart from the historical chronology.

We would like to stress at this point from the outset that the Hilbert space that we will construct in the course of the section is *just one from infinitely many inequivalent (kinematical) representations of the abstract algebra of operators*. On the other hand, as we will show in section I.3, it has many extremely natural and physically appealing properties and is at the moment the one that is most studied. However, one should not forget that there are many other (kinematical) representations which is a freedom that we may need to use exploit in later stages of the development of the theory. For some examples see section II.3.

I.2.1 The Space of Distributional Connections for Diffeomorphism Invariant Quantum Gauge Theories

In this section we will follow closely Velhinho [132]. For simplicity we stick to the analytic category. For generalization to the other categories discussed above, please refer to to the literature cited there. So in what follows, σ is an analytic, connected and orientable $D-$dimensional manifold which is locally compact (every point has a compact neighbourhood, automatic if σ is finite dimensional) and paracompact (the countable union of compact sets). Generalization to non-connected and non-orientable σ is straightforward.

I.2.1.1 The Label Set: Piecewise Analytic Paths

In all that follows we work with connection potentials, thus we assume that a fixed trivialization of the principal $G-$bundle has been chosen (upon passage to the gauge invariant sector nothing will depend on that choice any more).

Definition I.2.1
By $\mathcal{C}$ we denote the set of continuous, oriented, piecewise analytic, parameterized, compactly supported curves embedded into σ. That is, an element $c \in \mathcal{C}$ is given as a map

$$c : \; [0,1] \to \sigma; \; t \mapsto c(t) \tag{I.2.1. 1}$$

such that there is a finite natural number n and a partition $[0,1] = [t_0 = 0, t_1] \cup [t_1, t_2] \cup .. \cup [t_{n-1}, t_n = 1]$ and such that a) c is continuous at t_k, $k = 1,..,n-1$, b) real analytic in $[t_{k-1}, t_k]$, $k = 1,..,n-1$ and c) $c((t_{k-1}, t_k)$, $k = 1,..,n-1$ is an embedded one-dimensional submanifold of σ. Moreover, there is a compact subset of σ containing c.

Recall that a differentiable map $\phi : \; M_1 \to M_2$ between finite dimensional manifolds M_1, M_2 is called an immersion when ϕ has everywhere rank $\dim(M_1)$. An immersion need not be injective but when it is, it is called an embedding. For an embedding, the map $\phi : \; M_1 \to \phi(M_1)$ is a bijection and the manifold structure induced by ϕ on $\phi(M_1)$ is given by the atlas $\{\phi(U_I), \varphi_I \circ \phi^{-1}\}$ where $\{U_I, \varphi_I\}$ is an atlas of M_1. This differentiable structure need not be equivalent to the submanifold structure of $\phi(M_1)$ which is given by the atlas $\{V_J \cap \phi(M_1), \phi_J\}$ where $\{V_J, \phi_J\}$ is an atlas of M_2. When both differential structures are equivalent (diffeomorphic in the chosen differentiability category, say C^r, $r \in \mathbb{N} \cup \{\infty\} \cup \{\omega\}$ where ∞, ω denotes smooth and analytic respectively) the embedding is called regular. The above definition allows a curve to have self-intersections and self-overlappings so that it is only an immersion, but on the open intervals (t_{k-1}, t_k) a curve c is a regular embedding, in particular, it does not come arbitrarily close to itself.

Definition I.2.2
i) The beginning point, final point and range of a curve $c \in \mathcal{C}$ is defined, respectively, by

$$b(c) := c(0), \; f(c) := c(1), \; r(c) := c([0,1]) \tag{I.2.1. 2}$$

ii) Composition $\circ : \; \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ of composable curves $c_1, c_2 \in \mathcal{C}$ (those with $f(c_1) = b(c_2)$) and inversion $^{-1} : \; \mathcal{C} \to \mathcal{C}$ of $c \in \mathcal{C}$ are defined by

$$(c_1 \circ c_2)(t) \left\{ := \begin{array}{cc} c_1(2t) & t \in [0, \frac{1}{2}] \\ c_2(2t-1) & t \in [\frac{1}{2}, 1] \end{array} \right. , \; c^{-1}(t) := c(1-t) \tag{I.2.1. 3}$$

Notice that the operations (I.2.0. 17) do not equip $\mathcal{C}$ with the structure of a group for several reasons: First of all, not every two curves can be composed. Secondly, composition is notassociative because $(c_1 \circ c_2) \circ c_3, \; c_1 \circ (c_2) \circ c_3)$ differ by a reparametrization. Finally, the retraced curve $c \circ c^{-1}$ is not really just given by $b(c)$ so that c^{-1} is not the inverse of c and anyway there is no natural "identity" curve in $\mathcal{C}$.

Definition I.2.3
Two curves $c, c' \in \mathcal{C}$ are said to be equivalent, $c \sim c'$ if and only if
1) $b(c) = b(c')$, $f(c) = f(c')$ (identical boundaries) and
2) c' is identical with c up to a combination of a finite number of retracings and a reparameterization.

It is easy to see that $\sim$ defines an equivalence relation on $\mathcal{C}$ (reflexive: $c \sim c$, symmetric: $c \sim c' \Rightarrow c' \sim c$, transitive: $c \sim c',\ c' \sim c'' \Rightarrow c \sim c''$). The equivalence class of $c \in \mathcal{C}$ is denoted by p_c and the set of equivalence classes is denoted by $\mathcal{P}$. In order to distinguish the equivalence classes from their representative curves we wil refer to them as paths. As always, the dependence of $\mathcal{P}$ on σ will not be explicitly displayed. The second condition means that $c' = c_1' \circ \tilde{c}_1' \circ (\tilde{c}_1')^{-1} \circ .. \circ c_{n-1}' \circ \tilde{c}_{n-1}' \circ (\tilde{c}_{n-1}')^{-1} \circ c_n'$ for some finite natural number n and curves $c_k', \tilde{c}_l',\ k = 1,..,n,\ l = 1,..,n-1$ and that there exists a diffeomorphism $f:\ [0,1] \to [0,1]$ such that $c \circ f = c_1' \circ .. \circ c_n'$.

Definition I.2.3 has the following fibre bundle theoretic origin (see e.g. [139] and section III.2): Recall that a connection ω on a principal G bundle P maybe defined in terms of local connection potentials $A_I(x)$ over the chart U_I of an atlas $\{U_I, \varphi_I\}$ of σ which are the pull-backs to σ by local sections $s_I^\phi(x) := \phi_I(x, 1_G)$ of ω where $\phi_I:\ U_I \times G \to \pi^{-1}(U_I)$ denotes the system of local trivializations of P adapted to the U_I and π is the projection of P. The holonomy $h_{cI} := h_{cI}(1)$ of A_I along a curve in the domain of a chart U_I is uniquely defined by the differential equation

$$\dot{h}_{cI}(t) = h_{cI}(t) A_{Ia}(c(t)) \dot{c}^a(t);\ h_{cI}(0) = 1_G \qquad \text{(I.2.1. 4)}$$

and one may check that under a gauge transformation

$$A_I(x) \mapsto A_J(x) = -dh_{JI}(x) h_{JI}(x)^{-1} + \mathrm{ad}_{h_{JI}}(x)(A_I(x)) \qquad \text{(I.2.1. 5)}$$

the holonomy transforms as

$$h_{cI} \mapsto h_{cJ} = h_{JI}(b(c)) h_{cI} h_{JI}(f(c))^{-1} \qquad \text{(I.2.1. 6)}$$

Denote by $\mathcal{A}_P$ the space of smooth connections (abusing the notation by identifying the collection of potentials with the connection itself) over σ (the dependence on the bundle is explicitly displayed) and in what follows we will write $h_c(A)$ for the holonomy of A along c understood as an element of G which is possible once a trivialization has been fixed. We will denote by $A^g := -dgg^{-1} + \mathrm{ad}_g(A)$ a gauge transformed connection and have

$$h_c^g(A) := h_c(A^g) = g(b(c)) h_c(A) g(f(c))^{-1} \qquad \text{(I.2.1. 7)}$$

Besides these transformation properties, the holonomy has the following important algebraic properties:
1) $h_{c_1 \circ c_2}(A) = h_{c_1}(A) h_{c_2}(A)$,
2) $h_{c^{-1}}(A) = h_c(A)^{-1}$
as may be easily checked by using the differential equation (I.2.1. 4). Furthermore, one can verify that the differential equation (I.2.1. 4) is invariant under reparametrizations of c. These three properties guarantee that $h_c(A)$ does not depend on $c \in \mathcal{C}$ but only on the equivalence class $p_c \in \mathcal{C}$.

One might therefore also have given the following definition of equivalence of curves:

Definition I.2.4
Two curves $c, c' \in \mathcal{C}$ are said to be equivalent, $c \sim c'$ if and only if
1) $b(c) = b(c'),\ f(c) = f(c')$ (identical boundaries) and
2) $h_c(A) = h_{c'}(A)$ for all $A \in \mathcal{A}$.

In fact, defintions I.2.4) and I.2.3 are equivalent if G is compact and non-Abelean [117] since then every group element can be written as a commutator, that is, in the form $h = h_1 h_2 h_1^{-1} h_2^{-1}$ so that curves of the form $c_1 \circ c_2 \circ c_1^{-1}$ is not equivalent with c_2. In the Abelean case, definition I.2.4 is stronger than definition I.2.3. In what follows we will work with definition I.2.3.

Property 1) of definition I.2.1. 4 implies that the functions b, f can be extended to $\mathcal{C}$ by $b(p_c) := b(c), f(p_c) =: f(c)$, the right hand sides are independent of the representative. However, the function r can be extended only special elements which we will call edges.

Definition I.2.5
An edge $e \in \mathcal{P}$ is an equivalence class of a curve $c_e \in \mathcal{C}$ which is analytic in all of $[0, 1]$. *In this case* $r(e) := r(c_e)$.

For an entire analytic curve we may find an equivalent one which is not entire analytic and contains a retracing. However, we do not allow such representatives in the definition of $r(e)$.

It may be checked that $p_{c_1} \circ p_{c_2} := p_{c_1 \circ c_2}$ and $p_c^{-1} := p_{c^{-1}}$ are well-defined. The advantage of dealing with paths $\mathcal{P}$ rather than curves is that we now have almost a group structure since composition becomes associative and the path $p_c \circ p_c^{-1} = b(p_c)$ is trivial (stays at its beginning point). However, we still do not have a natural identity element in $\mathcal{P}$ and not all of its elements can be composed. The natural structure behind this is that of a *groupoid*. Let us recall the slightly more definition of a category.

Definition I.2.6
i)
A category $\mathcal{K}$ is a class (in general, more general than a set), the members of which are called objects $x, y, z, ..$, together with a collection $M(\mathcal{K})$ of sets $hom(x, y)$ for each ordered pair of objects (x, y), the members of which are called morphisms. Between the sets of morphisms there is defined a composition operation

$$\circ : \ hom(x, y) \times hom(y, z) \to hom(x, z); (f, g) \mapsto f \circ g \qquad \text{(I.2.1. 8)}$$

which satisfies the two following rules:
a) Associativity: $f \circ (g \circ h) = (f \circ g) \circ h$ for all $f \in hom(w, x), g \in hom(x, y), h \in hom(y, z)$,
b) Identities: For every $x \in \mathcal{K}$ there exists a unique element $id_x \in hom(x, x)$ such that for all $y \in \mathcal{K}$ we have $id_x \circ f = f$ for all $f \in hom(y, x)$ and $f \circ id_x = f$ for all $f \in hom(x, y)$.
ii)
A subcategory $\mathcal{K}' \subset \mathcal{K}$ is a category which contains a subclass of the class of objects in $\mathcal{K}$ and for each pair of objects (x, y) in $\mathcal{K}'$ we have for the set of morphisms $hom'(x, y) \subset hom(x, y)$.
iii)
A morphism $f \in hom(x, y)$ is called an isomorphism provided there exists $g \in hom(y, x)$ such that $f \circ g = id_y$, $g \circ f = id_x$.
iv)
If $\mathcal{K}_1$, $\mathcal{K}_2$ are categories with collections of sets of morphisms $M(\mathcal{K}_1)$, $M(\mathcal{K}_2)$ respectively, then a map $F : \ [\mathcal{K}_1, M(\mathcal{K}_1)] \to [\mathcal{K}_2, M(\mathcal{K}_2)]$ is called a covariant [contravariant] functor, also denoted by F_* $[F^*]$, provided that the algebraic structures are preserved, that is
1) $f \in hom(x, y) \ \Rightarrow \ F(f) \in hom(F(x), F(y))$ $[hom(F(y), F(x))]$
2) $F(f \circ g) = F(f) \circ F(g)$ $[F(g) \circ F(f)]$
3) $F(id_x) = id_{F(x)}$.
v)
A category in which every morphism is an isomorphism is called a groupoid.

This definition obviously applies to our situation with the following identifications:
Category: σ.
Objects: points $x \in \sigma$.
Morphisms: paths between points $hom(x, y) := \{p \in \mathcal{P}; b(p) = x, \ f(p) = y\}$. Obviously, every morphism is an isomorphism.
Collection of sets of morphisms : all paths $M(\sigma) = \mathcal{P}$
Composition: composition of paths $p_{c_1} \circ p_{c_2} = p_{c_1 \circ c_2}$
Identities: $\mathrm{id}_x = p \circ p^{-1}$ for any $p \in \mathcal{P}$ with $b(p) = x$.

We wil call this category σ the category of points and paths and denote it synonymously by $\mathcal{P}$ as well.
Subcategories: $\gamma \subset \mathcal{P}$ consisting of a subset of σ as the set of objects and for each two such objects x, y a subset $\hom'(x, y) \subset \hom(x, y)$.

It is clear that every path is a composition of edges, however, $\mathcal{P}$ is not freely generated by edges (free of algebraic relations among edges) because the composition $e \circ e'$ of two edges e, e' defined as the equivalence class of entire analytic curves $c_e, c_{e'}$ which are analytic continuations of each other defines a new edge e'' again. Notice that $\hom(x, y) \neq \emptyset$ for any $x, y \in \sigma$ because we have assumed that σ is connected, one says that $\mathcal{P}$ is connected. Moreover, $\hom(x, x)$ is actually a group with the identity element id_x being given by the trivial path in the equivalence class of the curve $c(t) = x,\ t \in [0, 1]$. The groups $\hom(x, x)$ are all isomorphic: Fix an arbitrary path $p_{xy} \in \hom(x, y)$, then $\hom(x, x) = p_{xy} \circ \hom(y, y) \circ p_{xy}^{-1}$.

Definition I.2.7
Fix once and for all $x_0 \in \sigma$. Then $\mathcal{Q} := hom(x_0, x_0)$ is called the hoop group in the literature.

The name "hoop" is an acronym for "holonomical equivalence class of a loop based at x_0". We use the word hoop to distinguish a hoop (a closed path) from its representative loop (a closed curve).

Lemma I.2.1
Fix once and for all a system of paths $p_x \in hom(x_0, x)$ with $p_{x_0} = id_{x_0}$. Then for any $p \in \mathcal{P}$ there is a unique $\alpha \in \mathcal{Q}$ such that

$$p = p_{b(p)}^{-1} \circ \alpha \circ p_{f(p)} \tag{I.2.1. 9}$$

The proof consists in solving equation (I.2.1. 9) for α.

Lemma I.2.2
Denote, for any subgroupoid $l \subset \mathcal{P}$ containing x_0 as an object, by $hom_l(x_0, x_0)$ the subgroup of $\mathcal{Q}$ consisting of hoops within γ.

Let $\mathcal{Q}'$ be any subgroup of $\mathcal{Q}$ and let $X \subset \sigma$ be any subset containing x_0. Then $l := \{p_x^{-1} \circ \alpha \circ p_y;\ x, y \in X,\ \alpha \in \mathcal{Q}'\}$ is a connected subgroupoid of $\mathcal{Q}$ (p_x the above fixed path system) and $\mathcal{Q}' = hom_l(x_0, x_0)$.

Proof of Lemma I.2.2:
i) l is a connected subgroupoid:
Given $p \in l$ there exist $x, y \in X,\ \alpha \in \mathcal{Q}'$ such that $p = p_x^{-1} \circ \alpha \circ p_y$. Thus $p^{-1} = p_y^{-1} \circ \alpha^{-1} \circ p_x \in l$ since $\mathcal{Q}'$ is a subgroup. Also given $p' = p_y^{-1} \circ \beta \circ p_z \in l$ we have $p \circ p' = p_x^{-1} \circ \alpha \circ \beta \circ p_z \in l$ since $\mathcal{Q}'$ is a subgroup. l is trivially connected since by construction every $x \in X$ is connected to $x_0 \in X$ through the path $p_x^{-1} \circ \alpha \circ p_y$ with $y = x_0, \alpha = \mathrm{id}_{x_0}$.
ii)
We have

$$\mathrm{hom}_l(x_0, x_0) = \{p \in \mathcal{Q};\ p \in l\} = \{p_{x_0}^{-1} \circ \alpha \circ p_{x_0};\ \alpha \in \mathcal{Q}'\} = \mathcal{Q}' \tag{I.2.1. 10}$$

since $p_{x_0} = \mathrm{id}_{x_0}$.
□

I.2.1.2 The Topology: Tychonov Topology

We have noticed above that for an element $A \in \mathcal{A}$ its holonomy $h_c(A)$ (understood as taking values in G, subject to a fixed trivialization) depends only on p_c. To express this we will use the notation

$$A(p_c) := h_c(A) \tag{I.2.1. 11}$$

It follows then that

$$A(p \circ p') = A(p)A(p'),\ A(p^{-1}) = A(p)^{-1} \tag{I.2.1. 12}$$

in other words, every $A \in \mathcal{A}_P$ defines a *groupoid morphism.*

Definition I.2.8
Hom$(\mathcal{P}, G)$ is the set of all (algebraic, no continuity assumptions) groupoid morphisms from the set of paths in σ into the gauge group.

What we have just shown is that $\mathcal{A}$ can be understood as a subset of $\mathrm{Hom}(\mathcal{P}, G)$ via the injection $H:\ \mathcal{A}_P \to \mathrm{Hom}(\mathcal{P}, G);\ A \mapsto H_A$ where $H_A(p) := A(p)$. That H is an injection ($H_A = H_{A'}$ implies $A = A'$) is the content of Giles' theorem [96] and can easily be understood from the fact that for a smooth connection $A \in \mathcal{A}$ we have for short curves $c_\epsilon:\ [0,1] \to \sigma;\ c_\epsilon(t) = c(\epsilon t),\ 0 < \epsilon < 1$ an expansion of the form $h_{c_\epsilon}(A) = 1_G + \epsilon \dot{c}^a(0)A_a(c(0)) + o(\epsilon^2)$ so that $(\frac{d}{d\epsilon})_{\epsilon=0} h_{c_\epsilon}(A) = \dot{c}^a(0)A_a(c(0))$, that is, by varying the curve c we can recover A from its holonomy.

We now show that $\mathcal{A}$ is certainly not all of $\mathrm{Hom}(\mathcal{P}, G)$, i.e. H is not a surjection, suggesting that $\mathrm{Hom}(\mathcal{P}, G)$ is a natural distributional extension of $\mathcal{A}$:
First of all, as we have said before, unless σ is three dimensional and $G = SU(2)$ the bundle P is not necessarily trivial and the classical spaces $\mathcal{A}$ are all different for different bundles. However, the space $\mathrm{Hom}(\mathcal{P}, G)$ depends only on σ and not on any P which means that it contains all possible classical spaces $\mathcal{A}$ at once and thus is much larger. Beyond this union of all the $\mathcal{A}$ it contains distributional elements, for instance the following: Let $f:\ S^2 \to G$ be any map, $x \in \sigma$ any point. Given a path p choose a representant c_p. The curve c_p can pass through x only a finite number of times, say N times, due to piecewise analyticity (see below). At the k-th passage denote by $n_k^\pm$ the direction of $\dot{c}_p(t)$ at x when it enters (leaves) x. Then define $H(p) := [f(-n_1^-)^{-1} f(n_1^+)]..[[f(-n_N^-)^{-1} f(n_N^+)]$ (for $N = 0$ defined to be 1_G). Notice that a retracing through x does not affect this formula because in that case $n_k^+ = -n_k^-$ and since we are taking only the direction of a tangent, also reparameterizations do not affect it. It follows that it depends only on paths rather than curves. It is easy to check that this defines an element of $\mathrm{Hom}(calP, G)$. It is not of the form $H_A,\ A \in \mathcal{A}_P$ because H has support only at x, it is distributional. More examples of distributional elements can be found in [106].

Having motivated the space $\mathrm{Hom}(\mathcal{P}, G)$ as a distributional extension of $\mathcal{A}_P$, the challenge is now to equip this so far only algebraically defined space with a topology. The reason is that, being distributional, it is a natural candidate for the support of a quantum field theory measure as we have stressed before but measure theory becomes most powerful in the context of topology. In order to define such a topology, projective techniques [58] suggest themselves. We begin quite general.

Definition I.2.9
i)
Let $\mathcal{L}$ be some abstract label (index) set. A partial order $\prec$ on $\mathcal{L}$ is a relation, i.e. a subset of $\mathcal{L} \times \mathcal{L}$, which is reflexive ($l \prec l$), symmetric ($l \prec l',\ l' \prec l \Rightarrow l = l'$) and transitive ($l \prec l',\ l' \prec l'' \Rightarrow l \prec l''$). Not all pairs of elements of $\mathcal{L}$ need to be in relation and if they are, $\mathcal{L}$ is said to be linearly ordered.
ii)
A partially ordered set $\mathcal{L}$ is said to be directed if for any $l, l' \in \mathcal{L}$ there exists $l'' \in \mathcal{L}$ such that $l, l' \prec l''$.
iii)
Let $\mathcal{L}$ be a partially ordered, directed index set. A projective family $(X_l, p_{l'l})_{l \prec l' \in \mathcal{L}}$ consists of sets X_l labelled by $\mathcal{L}$ together with surjective projections

$$p_{l'l}:\ X_{l'} \to X_l\ \forall\ l \prec l' \tag{I.2.1. 13}$$

satisfying the consistency condition

$$p_{l'l} \circ p_{l''l'} = p_{l''l} \ \forall \ l \prec l' \prec l'' \tag{I.2.1. 14}$$

iv)
The projective limit $\overline{X}$ of a projective family $(X_l, p_{l'l})$ is the subset of the direct product $X_\infty := \prod_{l \in \mathcal{L}} X_l$ defined by

$$\overline{X} := \{(x_l)_{l \in \mathcal{L}};\ p_{l'l}(x_{l'}) = x_l \ \forall \ l \prec l'\} \tag{I.2.1. 15}$$

The idea to use this definition for our goal to equip $\mathrm{Hom}(\mathcal{P}, G)$ with a topology is the following: We will readily see that $\mathrm{Hom}(\mathcal{P}, G)$ can be displayed as a projective limit. The compactness of the Hausdorff space G will be responsible for the fact that every X_l is compact and Hausdorff. Now on a direct product space (independent of the cardinality of the index set) in which each factor is compact and Hausdorff one can naturally define a topology, the so-called Tychonov topology, such that X_∞ is compact again. If we manage to show that $\overline{X}$ is closed in X_∞ then $\overline{X}$ will be compact and Hausdorff as well in the subspace topology (see, e.g. [140]). However, for compact Hausdorff spaces powerful measure theoretic theorems hold which will enable us to equip $\mathrm{Hom}(\mathcal{P}, G)$ with the structure of a $\sigma-$algebra and to develop measure theory thereon.

In order to apply definition I.2.9 then to our situation, we must decide on the label set $\mathcal{L}$ and the projective family.

Definition I.2.10 *i)*
A finite set of edges $\{e_1, .., e_n\}$ is said to be independent provided that the e_k intersect each other at most in the points $b(e_k), f(b_k)$. ii)
A finite set of edges $\{e_1, .., e_n\}$ is said to be algebraically independent provided none of the e_k is a finite composition of the $e_1, .., e_{k-1}, e_{k+1}, .., e_n$ and their inverses.
iii)
An independent set of edges $\{e_1, .., e_n\}$ defines an oriented graph γ by $\gamma := \cup_{k=1}^n r(e_k)$ where $r(e_k) \subset \gamma$ carries the arrow induced by e_k ($e \cup e' := p_{c_e \cup c_{e'}}$). From γ we can recover its set of edges $E(\gamma) = \{e_1, .., e_n\}$ as the maximal analytic segments of γ together with their orientations as well as set of vertices of γ as $V(\gamma) = \{b(e), f(e);\ e \in E(\gamma)\}$. Denote by Γ_0^ω the set all of all graphs.
iv)
Given a graph γ we denote by $l(\gamma) \subset \mathcal{P}$ the subgroupoid generated by γ with $V(\gamma)$ as the set of objects and with the $e \in E(\gamma)$ together with their inverses and finite compositions as the set of homomorphisms.

Notice that independence of sets of edges implies algebraic independence but not vice versa (consider independent e_1, e_2 with $f(e_1) = b(e_2)$ and define $e_1' = e_2,\ e_2' = e_1 \circ e_2$. Then e_1', e_2' is algebraically independent but not independent) and that $l(\gamma)$ is freely generated by the $e \in E(\gamma)$ due to their algebraic independence. Also, $l(\gamma)$ does not depend on the orientation of the graph since $e_1, .., e_n$ and $e_1^{s_1}, .., e_n^{s_n},\ s_k = \pm 1$ generate the same subgroupoid. The labels $\omega, 0$ in Γ_0^ω stand for "analytic" and "of compact support" respectively for obvious reasons.

The following theorem finally explains why it was important to stick with the analytic, compact category.

Theorem I.2.1
Let $\mathcal{L}$ be the set of all tame subgroupoids $l(\gamma)$ of $\mathcal{P}$, that is, those determined by graphs $\gamma \in \Gamma_0^\omega$. Then the relation $l \prec l'$ iff l is a subgroupoid of l' equips $\mathcal{L}$ with the structure of a partially ordered and directed set.

Proof of Theorem I.2.1:
Since l is a subgroupoid of l' iff all objects of l are objects of l' and all morphisms of l are morphisms of l' it is clear that $\prec$ defines a partial order. To see that $\mathcal{L}$ is directed consider any two graphs $\gamma, \gamma' \in \Gamma_0^\omega$ and consider $\gamma'' := \gamma \cup \gamma'$. We claim that γ'' has a finite number of edges again, that is, it is an element of Γ_0^ω. For this to be the case it is obviously sufficient to show that any two edges $e, e' \in \mathcal{P}$ can only have a finite number of isolated intersections or they are analytic extensions of each other. Clearly they are analytic extensions of each other if $e \cap e'$ is a common finite segment. Suppose then that $e \cap e'$ is an infinite discrete set of points. We may choose parameterizations of their representatives c, c' such that each of its component functions $f(t)^a := e'(t)^a - e(t)^a$ vanishes in at least a countably infinite number of points $t_m,\ m = 1, 2, ...$ We now show that for any function $f(t)$ which is real analytic in $[0,1]$ this implies $f = 0$. Since $[0,1]$ is compact there is an accumulation point $t_0 \in [0,1]$ of the t_m (here the compact support of the $c \in \mathcal{C}$ comes into play) and we may assume without loss of generality that t_m converges to t_0 and is strictly monotonous. Since f is analytic we can write the absolutely convergent Taylor series $f(t) = \sum_{n=0}^\infty f_n (t - t_0)^n$ (here analyticity comes into play). We show $f_n = 0$ by induction over $n = 0, 1, ...$ The induction start $f_0 = f(t_0) = \lim_{m \to \infty} f(t_m) = \lim_{m \to \infty} 0 = 0$ is clear. Suppose we have shown already that $f_0 = .. = f_n = 0$. Then $f(t) = f_{n+1}(t - t_0)^{n+1} + r_{n+1}(t)(t - t_0)^{n+2}$ where $r_{n+1}(t)$ is uniformly bounded in $[0,1]$. Thus $0 = f(t_m)/(t_m - t_0)^{n+1} = f_{n+1} + r_{n+1}(t_m)(t_m - t_0)$ for all m, hence $f_{n+1} = \lim_{m \to \infty}[f_{n+1} + r_{n+1}(t_m)(t_m - t_0)] = 0$.
$\Box$
Notice that the subgroupoids $l \in \mathcal{L}$ also conversely define a graph up to orientation through its edge generators.

Now that we have a partially ordered and directed index set $\mathcal{L}$ we must specify a projective family.

Definition I.2.11
For any $l \in \mathcal{L}$ define $X_l := Hom(l, G)$ the set of all homomorphisms from the subgroupoid l to G.

Notice that for $l = l(\gamma)$ any $x_l \in X_l$ is completely determined by the group elements $x_l(e),\ e \in E(\gamma)$ so that we have a bijection

$$\rho_\gamma :\ X_l \to G^{|E(\gamma)|};\ x_l \mapsto (x_l(e))_{e \in E(\gamma)} \tag{I.2.1. 16}$$

Since G^n for any finite n is a compact Hausdorff space (here compactness of G comes into play) in its natural manifold topology we can equip X_l with a compact Hausdorff topology through the identification (I.2.1. 16). This topology is independent of the choice of edge generators of l since any map $(e_1, .., e_n) \mapsto (e_{\pi(1)}^{s_1}, .., e_{\pi(n)}^{s_n})$ for any element $\pi \in S_n$ of the permutation group of n elements induces a homeomorphism (topological isomorphism) $G^n \to G^n$.

Next we must define the projections.

Definition I.2.12
For $l \prec l'$ define a projection by

$$p_{l'l} :\ X_{l'} \to X_l;\ x_{l'} \mapsto (x_{l'})_l \tag{I.2.1. 17}$$

restriction of the homomorphism $x_{l'}$ defined on the groupoid l' to its subgroupoid $l \prec l'$.

It is clear that the projection (I.2.1. 17) satisfies the consistency condition (I.2.1. 14) since for $l \prec l''$ we have $(x_{l''})_l = ((x_{l''})_{l'})_l$ for any intermediate $l \prec l' \prec l''$. Surjectivity is less obvious.

Lemma I.2.3
The projections $p_{l'l},\ l \prec l'$ are surjective, moreover, they are continuous.

Proof of Lemma I.2.3:
Let $l = l(\gamma) \prec l' = l(\gamma')$ be given. Since l is a subgroupoid of l' we may decompose any generator $e \in E(\gamma)$ in the form
$$e = \circ_{e' \in E(\gamma')} (e')^{s_{ee'}} \tag{I.2.1. 18}$$
where $s_{ee'} \in \{\pm 1, 0\}$. Notice that $|s_{ee'}| > 2$ is not allowed and that any e' appears at most once in (I.2.1. 18) because e is an edge (cannot overlap itself).
Surjectivity:
We must show that for any $x_l \in X_l$ there exists an $x_{l'} \in X_{l'}$ such that $p_{l'l}(x_{l'}) = x_l$. Since x_l is completely determined by $h_e := x_l(e) \in G,\ e \in E(\gamma)$ and $x_{l'}$ is completely determined by $h'_{e'} := x_{l'}(e') \in G,\ e' \in E(\gamma')$ and since h_e could be any value in G, what we have to show is that there exist group elements $h'_{e'} \in G,\ e' \in E(\gamma')$ such that for any group elements $h_e \in G,\ e \in E(\gamma)$ we have
$$h_e = \circ_{e' \in E(\gamma')} h_{e'}^{s_{e,e'}} \tag{I.2.1. 19}$$
However, since the $e \in E(\gamma)$ are disjoint up to their boundaries we have $s_{ee'} s_{e\tilde{e}'} = 0$ for any $e' \neq \tilde{e}'$ in $E(\gamma')$ so that we may specify some $e'(e) \in E(\gamma')$ for any $e \in E(\gamma)$ and the $e'(e)$ are disjoint up to their boundaries. Since also the $h'_{e'}$ can independently take any value we may choose $h'_{e'(e)} = h_e,\ h'_{e'} = 1_G$ for $e' \notin \{e'(e)\}_{e \in E(\gamma)}$.
Continuity:
Under the identification (I.2.1. 16) the projections are given as maps
$$p_{l'l}:\ G^{E(\gamma')} \to G^{E(\gamma)};\ (h'_{e'})_{e' \in E(\gamma')} \mapsto (\prod_{e' \in E(\gamma')} (h'_{e'})^{s_{ee'}})_{e \in E(\gamma)} \tag{I.2.1. 20}$$
By definition, a net $(h_k^\alpha)_{k=1}^n$ converges in G^n to $(h_k)_{k=1}^n$ if an only if every net $\lim_\alpha(h_k^\alpha) = h_k,\ k = 1,..,n$ individually converges (i.e., $(h_k^\alpha)_{AB} - (h_k)_{AB} \to 0$ for all matrix elements AB). Suppose then that $(h'^\alpha_{e'})_{e' \in E(\gamma')}$ converges to $(h'_{e'})_{e' \in E(\gamma')}$. By definition, in a Lie group inversion and finite multiplication are continuous operations. Therefore $(\prod_{e' \in E(\gamma')} (h'^\alpha_{e'})^{s_{ee'}})_{e \in E(\gamma)}$ converges to $(\prod_{e' \in E(\gamma')} (h'_{e'})^{s_{ee'}})_{e \in E(\gamma)}$ (as one can check also explicitly).
$\square$
We can now form the projective limit $\overline{X}$ of the X_l. In order to equip it with a topology we start by providing the direct product X_∞ with a topology. The natural topology on the direct product is the Tychonov topology.

Definition I.2.13
The Tychonov topology on the direct product $X_\infty = \prod_{l \in \mathcal{L}} X_l$ of topological spaces X_l is the weakest topology such that all the projections
$$p_l:\ X_\infty \to X_l;\ (x_{l'})_{l' \in \mathcal{L}} \mapsto x_l \tag{I.2.1. 21}$$
are continuous, that is, a net $x^\alpha = (x_l^\alpha)_{l \in \mathcal{L}}$ converges to $x = (x_l)_{l \in \mathcal{L}}$ iff $x_l^\alpha \to x_l$ for every $l \in \mathcal{L}$ pointwise (not necessarily uniformly) in $\mathcal{L}$.

We then have the following non-trivial result.

Theorem I.2.2 (Tychonov)
Let $\mathcal{L}$ be an index set of arbitrary cardinality and suppose that for each $l \in \mathcal{L}$ a compact topological space X_l is given. Then the direct product space $X_\infty = \prod_{l \in \mathcal{L}} X_l$ is a compact toplogical space in the Tychonov topology.

An elegant proof of this theorem in terms of universal nets is given in section III.3 where also other relevant results from general topology inluding proofs can be found.

Since $\overline{X} \subset X_\infty$ we may equip it with the subspace topology, that is, the open sets of $\overline{X}$ are the sets $U \cap \overline{X}$ where $U \subset X_\infty$ is any open set in X_∞.

Lemma I.2.4
The projective limit $\overline{X}$ is a closed subset of X_∞.

Proof of Lemma I.2.4:
Let $(x^\alpha) := ((x_l^\alpha)_{l\in\mathcal{L}})$ be a convergent net in X_∞ such that $x^\alpha := (x_l^\alpha)_{l\in\mathcal{L}} \in \overline{X}$ for any α. We must show that the limit point $x = (x_l)_{l\in\mathcal{L}}$ lies in $\overline{X}$. By lemma I.2.3, the projections $p_{l'l} : X_{l'} \to X_l$ are continuous, therefore

$$p_{l'l}(x_{l'}) = \lim_\alpha p_{l'l}(x_{l'}^\alpha) = \lim_\alpha x_l^\alpha = x_l \qquad \text{(I.2.1. 22)}$$

where the second equality follows from $x^\alpha \in \overline{X}$. Thus, the point $x \in X_\infty$ qualifies as a point in $\overline{X}$.
□
Since closed subspaces of compact spaces are compact in the subspace topology (see section III.3) we conclude that $\overline{X}$ is compact in the subspace topology induced by X_∞.

Lemma I.2.5
Both $X_\infty, \overline{X}$ are Hausdorff spaces.

Proof of Lemma I.2.5:
By assumtion, G is a Hausdorff topological group. Thus G^n for any finite n is a Hausdorff topological group as well and since X_l is topologically identified with some G^n via I.2.1. 16 we see that X_l is a topological Hausdorff space for any $l \in \mathcal{L}$. Let now $x \neq x'$ be points in X_∞. Thus, there is at least one $l_0 \in \mathcal{L}$ such that $x_{l_0} \neq x'_{l_0}$. Since X_{l_0} is Hausdorff we find disjoint open neighbourhoods $U_{l_0}, U'_{l_0} \subset X_{l_0}$ of x_{l_0}, x'_{l_0} respectively. Let $U := p_{l_0}^{-1}(U_{l_0}), U' := p_{l_0}^{-1}(U'_{l_0})$. Since the topology of X_∞ is generated by the continuous functions $p_l : X_\infty \to X_l$ from the topology of the X_l, it follows that U, U' are open in X_∞. Moreover, U, U' are obviously neighbourhoods of x, x' respectively since $p_l(U) = X_l = p_l(U')$ for any $l \neq l_0$. Finally, $U \cap U' = \emptyset$ since $p_{l_0}(U \cap U') = U_{l_0} \cap U'_{l_0} = \emptyset$ so that U, U' are disjoint open neighbourhoods of $x \neq x'$ and thus X_∞ is Hausdorff.

Finally, to see that $\overline{X}$ is Hausdorff, let $x \neq x'$ be points in $\overline{X}$, then we find respective disjoint open neighbourhoods U, U' in X_∞ whence $U \cap \overline{X}, U' \cap \overline{X}$ are disjoint open neighbourhoods in $\overline{X}$ by definition of the subspace topology.
□
Let us collect these results in the following theorem.

Theorem I.2.3
The projective limit $\overline{X}$ of the spaces $X_l = Hom(l, G)$, $l \in \mathcal{L}$ where $\mathcal{L}$ denotes the set of all tame subgroupoids of $\mathcal{P}$ is a compact Hausdorff space in the induced Tychonov topology whenever G is a compact Hausdorff topological group.

The purpose of our efforts was to equip $\text{Hom}(\mathcal{P}, G)$ with a topology. Theorem I.2.2 now enables us to do this provided we manage to identify $\text{Hom}(\mathcal{P}, G)$ with the projective limit $\overline{X}$ via a suitable bijection. Now an elementary exercise is that any point of $\text{Hom}(\mathcal{P}, G)$ defines a point in $\overline{X}$ if we define $x_l := H_{|l}$ since the projections $p_{l'l}$ encode the algebraic relations that are induced by asking that H be a homomorphism. That this map is actually a bijection is the content of the following theorem.

Theorem I.2.4
The map
$$\Phi : \; Hom(\mathcal{P}, G) \to \overline{X}; \; H \mapsto (H_{|l})_{l\in\mathcal{L}} \tag{I.2.1. 23}$$
is a bijection.

Proof of theorem I.2.3:
Injectivity:
Suppose that $\Phi(H) = \Phi(H')$, in other words, $H_{|l} = H'_{|l}$ for any $l \in \mathcal{L}$. Thus, if $l = l(\gamma)$ we have $H(e) = H'(e)$ for any $e \in E(\gamma)$. Since l is arbitrary we find $H(p) = H'(p)$ for any $p \in \mathcal{P}$, that is, $H = H'$.
Surjectivity:
Suppose we are given some $x = (x_l)_{l\in\mathcal{L}} \in \overline{X}$. We must find $H_x \in \mathrm{Hom}(\mathcal{P}, G)$ such that $\Phi(H_x) = x$. Let $p \in \mathcal{P}$ be any path, then we can always find a graph γ_p such that $p \in l := l(\gamma_p)$. We may then define
$$H_x(p) := x_{l(\gamma_p)}(p) \tag{I.2.1. 24}$$
Of course, the map $p \mapsto \gamma_p$ is one to many and therefore the definition (I.2.1. 24) seems to be ill-defined. We now show that this is not the case, i.e., (I.2.1. 24) does not depend on the choice of γ_p. Thus, let γ'_p be any other graph such that $p \in l' := l(\gamma'_p)$. Since $\mathcal{L}$ is directed we find l'' with $l, l' \prec l''$. But then by the definition of a point x in the projective limit
$$x_l(p) = [p_{l''l}(x_{l''})](p) = (x_{l''})_{|l}(p) \equiv x_{l''}(p) \equiv (x_{l''})_{|l'}(p) = [p_{l''l}(x_{l''})](p) = x_{l'}(p) \tag{I.2.1. 25}$$
It remains to check that H_x is indeed a homomorphism. We have for any $p, p', p \circ p' \in l$ with $f(p) = b(p')$
$$H_x(p^{-1}) = x_l(p^{-1}) = (x_l(p))^{-1} = H_x(p)^{-1} \text{ and } H_x(p \circ p') = x_l(p \circ p') = x_l(p)x_l(p') = H_x(p)H_x(p') \tag{I.2.1. 26}$$
since $x_l \in \mathrm{Hom}(l, G)$.
$\square$

Definition I.2.14 *The space $\overline{\mathcal{A}} := Hom(\mathcal{P}, G)$ of homomorphisms from the set of piecewise analytical paths into the compact Hausdorff topological group G, identified set-theoretically and topologically via (I.2.1. 23) with the projective limit $\overline{X}$ of the spaces $X_l = Hom(l, G)$, where $l \in \mathcal{L}$ runs through the tame subgroupoids of $\mathcal{P}$, is called the space of distributional connections over σ. In the induced Tychonov topology inherited from $\overline{X}$ it is a compact Hausdorff space.*

Once again it is obvious that the space of distributions $\overline{\mathcal{A}}$ does not carry any sign anymore of the bundle P, it depends only on the base manifold σ via the set of embedded paths $\mathcal{P}$.

I.2.1.3 Gauge Invariance: Distributional Gauge Transformations

The space $\overline{\mathcal{A}}$ contains connections (from now on considered as morphisms $\mathcal{P} \to G$) which are nowhere continuous as we will see later on and these turn out to be measure-theoretically much more important than the smooth ones contained in $\mathcal{A}$. Therefore it is motivated to generalize also the space of smooth gauge transformations $\mathcal{G} := C^\infty(\sigma, G)$ to the space of *all* functions
$$\overline{\mathcal{G}} := \mathrm{Fun}(\sigma, G) \tag{I.2.1. 27}$$
with no restrictions (e.g. continuity). It is clear that $g \in \mathcal{G}$ may be thought of as the net $(g(x))_{x\in\sigma}$ and thus $\mathcal{G}$ is just the continuous infinite direct product $\mathcal{G} = \prod_{x\in\sigma} G$.

The transformation property of $\mathcal{A}$ under $\mathcal{G}$ (I.2.1. 7) can be understood as an action $\lambda : \mathcal{G}\times\mathcal{A}\to\mathcal{A};\ (g,A)\to A^g := \lambda_g(A) := \lambda(g,A)$ where $A^g(p) := g(b(p))A(p)g(f(p))^{-1}$ for any $p\in\mathcal{P}$ which we may simply lift to $\overline{\mathcal{A}},\overline{\mathcal{G}}$ as

$$\begin{aligned}\lambda &: \overline{\mathcal{G}}\times\overline{\mathcal{A}}\to\overline{\mathcal{A}};\ (g,A)\to A^g := \lambda_g(A) := \lambda(g,A) \text{ where}\\ A^g(p) &:= g(b(p))A(p)g(f(p))^{-1}\ \forall\, p\in\mathcal{P}\end{aligned} \tag{I.2.1. 28}$$

Notice that this is really an action, i.e. A^g really is an element of $\overline{\mathcal{A}} = \text{Hom}(\mathcal{P},G)$, that is, it satisfies the homomorphism property

$$\begin{aligned}A^g(p^{-1}) &= g(b(p^{-1}))A(p^{-1})g(f(p^{-1}))^{-1} = g(f(p))A(p)^{-1}g(b(p))^{-1} = (A^g(p))^1\\ A^g(p)A^g(p') &= [g(b(p))A(p)g(f(p))^{-1}][g(b(p'))A(p')g(f(p'))^{-1}] = g(b(p))A(p)A(p')g(f(p'))^{-1}\\ &= g(b(p\circ p'))A(p\circ p')g(f(p\circ p'))^{-1} = A^g(p\circ p')\end{aligned} \tag{I.2.1. 29}$$

because $f(p)=b(p'), b(p)=b(p\circ p'), f(p')=f(p\circ p')$. The action (I.2.1. 28) is also continuous on $\overline{\mathcal{A}}$, that is, for any $g\in\overline{\mathcal{G}}$ the map $\lambda_g : \overline{\mathcal{A}}\to\overline{\mathcal{A}}$ is continuous. To see this, let (A^α) be a net in $\overline{\mathcal{A}}$ converging to $A\in\overline{\mathcal{A}}$. Then $\lim_\alpha \lambda_g(A_\alpha)) = \lambda_g(A)$ if and only if $\lim_\alpha p_l(\lambda_g(A_\alpha))) = p_l(\lambda_g(A))$ for any $l\in\mathcal{L}$. Identifying $\overline{\mathcal{A}}_{|l}$ with some G^n via (I.2.1. 16) and using the bijection (I.2.1. 23) we have for any $p\in l$

$$\begin{aligned}[p_l(\lambda_g(A^\alpha))](l) &= [(\lambda_g(A^\alpha))_{|l}](p) = [\lambda_g(A^\alpha)](p) = g(b(p))A^\alpha(p)g(f(p))^{-1}\\ &= g(b(p))[p_l(A^\alpha)](p)g(f(p))^{-1}\end{aligned} \tag{I.2.1. 30}$$

Since group multiplication and inversion are continuous in G^n we easily get $\lim_\alpha[p_l(\lambda_g(A^\alpha))](l) = [p_l(\lambda_g(A))](l)$ for any $p\in l$, that is, $\lim_\alpha p_l(\lambda_g(A^\alpha)) = p_l(\lambda_g(A))$, thus λ_g is continuous for any $g\in\overline{\mathcal{G}}$.

Since $\overline{\mathcal{A}}$ is a compact Hausdorff space and λ is a continuous group action on $\overline{\mathcal{A}}$ it then follows immediately from abstract results (see section III.3) that the quotient space

$$\overline{\mathcal{A}}/\overline{\mathcal{G}} := \{[A];\ A\in\overline{\mathcal{A}}\} \text{ where } [A] := \{A^g;\ g\in\overline{\mathcal{G}}\} \tag{I.2.1. 31}$$

is a compact Hausdorff space in the quotient topology. The quotient topology on the quotient $\overline{\mathcal{A}}/\overline{\mathcal{G}}$ is defined as follows: The open sets in $\overline{\mathcal{A}}/\overline{\mathcal{G}}$ are precisely those whose preimages under the quotient map

$$[] :\ \overline{\mathcal{A}}\to\overline{\mathcal{A}}/\overline{\mathcal{G}};\ A\mapsto[A] \tag{I.2.1. 32}$$

are open in $\overline{\mathcal{A}}$, that is, the quotient topology is generated by asking that the quotient map be continuous.

Now as $\overline{\mathcal{G}}$ is a continuous direct product of the compact Hausdorff spaces G it is a compact Hausdorff space in the Tychonov topology by the theorems proved in section I.2.1.2. More explicitly, the projective construction of $\overline{\mathcal{G}}$ proceeds as follows: Given $l\in\mathcal{L}$ with $l=l(\gamma)$ we define $\overline{\mathcal{G}}_l := \prod_{v\in V(\gamma)} G$ and extend the surjective projection $p_l :\ \overline{\mathcal{A}}\to\overline{\mathcal{A}}_l;\ A\mapsto A_{|l}$ to $p_l :\ \overline{\mathcal{G}}\to\overline{\mathcal{G}}_l;\ g\mapsto g_{|l}$ and for $l\prec l'$ the surjective projection $p_{l'l} :\ \overline{\mathcal{A}}_{l'}\to\overline{\mathcal{A}}_l;\ A_{l'}\mapsto(A_{l'})_{|l}$ to $p_{l'l} :\ \overline{\mathcal{G}}_{l'}\to\overline{\mathcal{G}}_l;\ g_{l'}\mapsto(g_{l'})_{|l}$. These projections are obviously surjective again because $\overline{\mathcal{G}}$ is actually a direct product of copies of G, one for every $x\in\sigma$.

Notice that the projective limit $\overline{\mathcal{G}} = \{(g_l)_{l\in\mathcal{L}};\ p_{l'l}(g_{l'}) = g_l\}$ is a group since $p_{l'l}(g_{l'}g'_{l'}) = g_l g'_l = p_{l'l}(g_{l'})p_{l'l}(g'_{l'})$ and $p_{l'l}((g^{-1})_{l'}) = (g^{-1})_l = g_l^{-1} = p_{l'l}(g_{l'})^{-1}$ so that actually the $p_{l'l}$ are surjective group homomorphisms. Since the $\overline{\mathcal{G}}_l$ are compact Hausdorff topological groups it follows that $\overline{\mathcal{G}}$ is also a compact Hausdorff topological group.

Summarizing: $\overline{\mathcal{A}}/\overline{\mathcal{G}}$ is the quotient of two projective limits both of which are compact Hausdorff spaces.

On the other hand observe that for $l \prec l'$ we have

$$p_{l'l}(\lambda_{g_{l'}}(A_{l'})) = \lambda_{g_l}(A_l) \tag{I.2.1. 33}$$

for any $A \in \overline{\mathcal{A}}, g \in \overline{\mathcal{G}}$, one says the group action λ is *equivariant*. Consider then the quotients

$$[\overline{\mathcal{A}}_l]_l := \overline{\mathcal{A}}_l/\overline{\mathcal{G}}_l := \{[A_l]_l;\ A_l \in \overline{\mathcal{A}}_l\} \text{ where } []_l:\ \overline{\mathcal{A}}_l \to \overline{\mathcal{A}}_l/\overline{\mathcal{G}}_l;\ A_l \mapsto [A_l]_l := \{\lambda_{g_l}(A_l);\ g_l \in \overline{\mathcal{G}}_l\} \tag{I.2.1. 34}$$

Due to the equivariance property for $l \prec l'$

$$p_{l'l}([A_{l'}]_{l'}) = \{p_{l'l}(\lambda_{g_{l'}}(A_{l'});\ g_{l'} \in \overline{\mathcal{G}}_{l'}\} = \{\lambda_{g_l}(A_l);\ g_l \in \overline{\mathcal{G}}_l\} = [A_l]_l \tag{I.2.1. 35}$$

since the projections $p_{l'l}:\ \overline{\mathcal{G}}_{l'} \to \overline{\mathcal{G}}_l$ are surjective. Now $\overline{\mathcal{A}}_l$ is a compact Hausdorff space and λ a continuous group action og $\overline{\mathcal{G}}_l$ thereon, thus $[\overline{\mathcal{A}}_l]_l$ is a compact Hausdorff space in the quotient topology induced by $[]_l$. By the results proved in section I.2.1.2 we find that the projective limit of these quotients, denoted by $\overline{\mathcal{A}/\mathcal{G}}$, is again a compact Hausdorff space in the induced Tychonov topology.

We therefore have two compact Hausdorff spaces associated with gauge invariance, on the one hand the quotient of projective limits $\overline{\mathcal{A}}/\overline{\mathcal{G}}$ and on the other hand the projective limit of the quotients $\overline{\mathcal{A}/\mathcal{G}}$. The question arises what the relation between the spaces $\overline{\mathcal{A}}/\overline{\mathcal{G}}, \overline{\mathcal{A}/\mathcal{G}}$ is. In what follows we will show by purely algebraic and topological methods (without using C^* algebra techniques) that they are homeomorphic.

We begin by giving a characterization of $\overline{\mathcal{A}/\mathcal{G}}$ similar to the characterization of $\overline{\mathcal{A}}$ as $\text{Hom}(\mathcal{P}, G)$.

Definition I.2.15
Let, as in definition I.2.7 a point $x_0 \in \sigma$ be fixed once and for all and denote by $\mathcal{Q} := hom(x_0, x_0)$ the hoop group of σ.
i)
A finite set $\{\alpha_1, .., \alpha_n\}$ of hoops is said to be independent if any α_k contains an edge that is traversed precisely once and that is intersected by any α_l, $l \neq k$ in at most a finite number of points. ii)
An independent set of hoops $\{\alpha_1, .., \alpha_n\}$ defines an unoriented, closed graph $\check{\gamma}$ by $\check{\gamma} := \cup_{k=1}^n r(\alpha_k)$ ($\alpha \cup \alpha' := p_{c_\alpha \cup c_{\alpha'}}$) up to x_0. Here closed up to x_0 means that every vertex is at least bivalent except, possibly for the vertex x_0. From an oriented graph γ we can recover one set $H(\gamma) = \{\beta_1, .., \beta_n\}$ of independent hoops generating the fundamental group $\pi_1(\gamma)$ of γ (although not a canonical one whence possibly $\{\alpha_k\} \neq \{\beta_k\}$ but the number n is identical for both sets) as well as the set of vertices of γ as $V(\gamma) = \{b(e), f(e);\ e \in E(\gamma)\}$. We fix once and for all generators of $\pi_1(\gamma)$ for every oriented graph γ.
iii)
Given a graph γ we denote by $s(\gamma) \subset \mathcal{Q}$ the (so-called tame) subgroup generated by the generators of $\pi_1(\gamma)$, that is, $s(\gamma) = \pi_1(\gamma)$.

We now have an analogue of theorem I.2.1

Theorem I.2.5
Let $\mathcal{S}$ be the set all tame subgroups $s(\gamma)$ of $\mathcal{Q}$, that is, those freely generated by graphs $\gamma \in \Gamma_0^\omega$. Then the relation $s \prec s'$ iff s is a subgroup of s' equips $\mathcal{Q}$ with the structure of a partially ordered and directed set.

Let now $Y_s := \mathrm{Hom}(s, G)$. As with $X_l = \mathrm{Hom}(l, G)$ we can identify Y_s with some G^n displaying it as a compact Hausdorff space. Likewise we have surjective projections for $s \prec s'$ given by the restriction map, $p_{s's}: Y_{s'} \to Y_s;\ x_{s'} \mapsto (x_{s'})_{|s}$ which satisfy the consistency condition $p_{s's} \circ p_{s''s'} = p_{s''s}$ for any $s \prec s' \prec s''$. We therefore can form the direct product $Y_\infty = \prod_{s\in\mathcal{S}} Y_s$ and its projective limit subset

$$\overline{Y} = \{y = (y_s)_{s\in\mathcal{S}};\ p_{s's}(ys') = y_s\ \forall\ s \prec s'\} \tag{I.2.1. 36}$$

which in the Tychonov topology induced from Y_∞ is a compact Hausdorff space. Repeating step by step the proof of theorem I.2.3) we find that the map

$$\Phi:\ \mathrm{Hom}(\mathcal{Q}, G) \to \overline{Y};\ H \mapsto (H_{|s})_{s\in\mathcal{S}} \tag{I.2.1. 37}$$

is a bijection so that we can identify $\mathrm{Hom}(\mathcal{Q}, G)$ with $\overline{Y}$ and equip it with the topology of $\overline{Y}$ (open sets of $\mathrm{Hom}(\mathcal{Q}, G)$ are the sets $\Phi^{-1}(U)$ where U is open in $\overline{Y}$). This topology is the weakest one so that all the projections $p_s:\ \overline{Y} \to Y_s;\ y \mapsto y_s$ are continuous.

The action λ of $\overline{\mathcal{G}}$ on $\overline{\mathcal{A}} = \overline{X}$ reduces on $\overline{Y}$ to

$$\lambda:\ \overline{\mathcal{G}} \times \overline{Y} \to \overline{Y};\ (g, y) \mapsto \lambda(g, y) = \lambda_g(y) = \mathrm{Ad}_g(y);\ [\mathrm{Ad}_g(y)]_s = \mathrm{Ad}_{g(x_0)}(y_s) \tag{I.2.1. 38}$$

where for $\alpha \in s$ we have $[\mathrm{Ad}_{g(x_0)}(y_s)](\alpha) = \mathrm{Ad}_{g(x_0)}(y(\alpha))$ and $\mathrm{Ad}:\ G \times G \to G;\ (g, h) \mapsto ghg^{-1}$ is the adjoint action of G on itself. In other words, $(\lambda_{\overline{\mathcal{G}}})|\overline{Y} = \mathrm{Ad}_G$ where G can be identified with the restriction of $\overline{\mathcal{G}}$ to x_0. Clearly Ad acts continuously on $\overline{Y}$.

Consider then the quotient space $\mathrm{Hom}(\mathcal{Q}, G)/G$ (notice that we mod out by G and not $\overline{\mathcal{G}}$!) which by the results obtained in the previous section is a compact Hausdorff space in the quotient topology. Now the action Ad on $\overline{Y}$ is completely independent of the label s, that is

$$\mathrm{Ad}_g \circ p_{s's} = p_{s's} \circ \mathrm{Ad}_g \tag{I.2.1. 39}$$

so that the points in $\overline{Y}/G$ are given by the equivalence classes

$$(y) := \{\mathrm{Ad}_g(y);\ g \in G\} = \{(\mathrm{Ad}_g(y_s))_{s\in\mathcal{S}};\ g \in G\} = ((y_s)_s)_{s\in\mathcal{S}} \tag{I.2.1. 40}$$

where $()_s:\ Y_s \to (Y_s)_s\ y_s \mapsto (y_s)_s = \{\mathrm{Ad}_g(y_s);\ g \in G\}$ denotes the quotient map in Y_s. It follows that $\mathrm{Hom}(\mathcal{Q}, G)/G$ is the projective limit of the $(Y_s)_s$. On the other hand, consider the quotients $[X_l]_l$ discussed above. If $l' = l(\gamma')$ and γ' is not a closed graph then by the action of $\overline{\mathcal{G}}$ on $X_{l'}$ we get $[X_{l'}]_{l'} = [X_l]_l$ where $l = l(\gamma)$ and γ is the closed graph obtained from γ' by deleting its open edges (monovalent vertices). Next, if $x_0 \notin \gamma$ then we add a path to γ connecting any of its points to x_0 without intersecting γ otherwise and obtain a third graph γ'' where again $[X_{l''}]_{l''} = [X_l]_l$ with $l'' = l(\gamma^d prime)$ due to quotienting by the action of the gauge group. But now γ'' is a closed graph up to x_0. Thus we see that the projective limit of the $[X_l]_l,\ l \in \mathcal{L}$ and of the $[Y_s]_l,\ s \in \mathcal{S}$ coincides, in other words we have the identity

$$\overline{\mathcal{A}/\mathcal{G}} = \mathrm{Hom}(\mathcal{Q}, G)/G \tag{I.2.1. 41}$$

Our proof of the existence of a homeomorphism between $\overline{\mathcal{A}}/\overline{\mathcal{G}}$ and $\overline{\mathcal{A}/\mathcal{G}}$ will be based on the identity (I.2.1. 41) and the fact that $\overline{\mathcal{A}} = \mathrm{Hom}(\mathcal{P}, G)$. We will break this proof into several lemmas.

Fix once and for all a system of edges

$$\mathcal{E} := \{e_x \in \mathrm{Hom}(x_0, x);\ x \in \sigma\} \tag{I.2.1. 42}$$

where e_{x_0} is the trivial hoop based at x_0. Let $\overline{\mathcal{G}}_{x_0} := \{g \in \overline{\mathcal{G}};\ g(x_0) = 1_G\}$ be the subset of all gauge transformtions that are the identity at x_0 and consider the following map

$$\begin{aligned} f_{\mathcal{E}} &: \mathrm{Hom}(\mathcal{P}, G) \to \mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0};\ A \mapsto (B, h) \text{ where} \\ B(\alpha) &: = A(\alpha)\ \forall\ \alpha \in \mathcal{Q} \text{ and } h(x) := A(e_x)\ \forall x \in \sigma \end{aligned} \tag{I.2.1. 43}$$

Clearly $g(x_0) = A(e_{x_0}) = 1_G$. From the known action λ of $\overline{\mathcal{G}}$ on $\overline{\mathcal{A}}$ we induce the following action of $\overline{\mathcal{G}}$ on $\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0}$

$$\begin{aligned} \lambda' &: \overline{\mathcal{G}} \times (\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0}) \to (\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0});\ (g, (B, h)) \mapsto (B^g, h^g) = \lambda'_g(B, h) \\ &\text{where} B^g(\alpha) = \mathrm{Ad}_{g(x_0)}(B(\alpha)); \forall\ \alpha \in \mathcal{Q} \text{ and } h^g(x) = g(x_0)h(x)g(x)^{-1}\ \forall\ x \in \sigma \end{aligned} \tag{I.2.1. 44}$$

The action (I.2.1. 44) evidently splits into a $G-$action by Ad on $\mathrm{Hom}(\mathcal{Q}, G)$ (with $G \equiv \overline{\mathcal{G}}_{|x_0}$) as already observed above and a $\overline{\mathcal{G}}-$action on $\overline{\mathcal{G}}_{x_0}$ (indeed $h^g(x_0) = 1_G$)).

Theorem I.2.6
For any choice of $\mathcal{E}$ the map $f_{\mathcal{E}}$ in (I.2.1. 43) is a homeomorphism which is $\lambda-$equivariant, that is,

$$f_{\mathcal{E}} \circ \lambda = \lambda' \circ f_{\mathcal{E}} \tag{I.2.1. 45}$$

Proof of Theorem I.2.6:
Bijection:
The idea is to construct explicitly the inverse $f_{\mathcal{E}}^{-1}$. The ansatz is of course, that given any $p \in \mathcal{P}$ we can construct a hoop based at x_0 by using $\mathcal{E}$, namely $\alpha_p := e_{b(p)} \circ p \circ e_{f(p)}^{-1}$, which we can use in order to evaluate a given $B \in \mathrm{Hom}(\mathcal{Q}, G)$. Since we want that $A^g(p) = g(b(p))A(p)g(f(p))^{-1}$ we see that given $h \in \overline{\mathcal{G}}_{x_0}$ the only possibility is

$$\begin{aligned} f_{\mathcal{E}}^{-1} &: \mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0} \to \mathrm{Hom}(\mathcal{P}, G);\ (B, h) \mapsto A \text{ where} \\ A(p) &:= h(b(p))^{-1} B(e_{b(p)} \circ p \circ e_{f(p)}^{-1}) h(f(p)) \end{aligned} \tag{I.2.1. 46}$$

One can verify explicitly that this is the inverse of (I.2.1. 43).
Equivariance:
Trivial by construction.
Continuity:
By definition of the topology on the spaces $\mathrm{Hom}(\mathcal{P}, G)$, $\mathrm{Hom}(\mathcal{Q}, G)$, $\overline{\mathcal{G}}$ respectively, a corresponding net $(A^\alpha), (B^\alpha), (g^\alpha)$ converges to A, B, g iff the nets $(A_l^\alpha) = (p_l(A^\alpha)), (B_s^\alpha) = (p_s(B^\alpha)), (g_x^\alpha) = (p_x(g^\alpha))$ converge to $A_l = p_l(A), B_s = p_s(B), g_x = p_x(g)$ where $g_x = g(x)$ for all $l \in \mathcal{L}, s \in \mathcal{S}, x \in \sigma$.

Continuity of $f_{\mathcal{E}}$ then means that $(p_s \times p_x) \circ f_{\mathcal{E}}$ is continuous for all $s \in \mathcal{S}, x \in \sigma$ while continuity of $f_{\mathcal{E}}^{-1}$ means that $p_l \circ f_{\mathcal{E}}^{-1}$ is continuous for all $l \in \mathcal{L}$. Recalling the map (I.2.1. 16) it is easy to see that

$$p_x \circ f_{\mathcal{E}} = \rho_{e_x} \circ p_{l(e_x)} \tag{I.2.1. 47}$$

and since the ρ_γ are by definition continuous we easily get continuity of $p_x \circ f_{\mathcal{E}}$ as the composition of two continuous maps.

To establish the continuity of $p_s \circ f_{\mathcal{E}}, p_l \circ f_{\mathcal{E}}^{-1}$ requires more work.

Lemma I.2.6
i)

For all $s \in \mathcal{S}$ there exists a connected subgroupoid $l \in \mathcal{L}$ such that s is a subgroup of l, i.e. $s \prec l$ ($s \in \mathcal{L}$ in particular). The projection

$$p_{ls}:\; X_l \to Y_s;\; x_l \mapsto (x_l)_{|s} \tag{I.2.1. 48}$$

is continuous and satisfies $p_s \circ f_{\mathcal{E}} = p_{ls} \circ p_l$ for any choice of $\mathcal{E}$.
ii)
For any $l \in \mathcal{L}$ there exists $s \in \mathcal{S}$ and a conncted subgroupoid $l' \in \mathcal{L}$ such that with $l = l(\gamma), l' = l(\gamma')$ we have $V(\gamma') = V(\gamma) \cup \{x_0\}$, moreover $l \prec l'$ and $hom_{l'}(x_0, x_0) = s$. Let $\overline{\mathcal{G}}_{x_0}(l') := Fun(V(\gamma'), G) \cap \overline{\mathcal{G}}_{x_0}$ and let $\pi_{l'}:\; \overline{\mathcal{G}}_{x_0} \to \overline{\mathcal{G}}_{x_0}(l')$ be the restriction map. The projection $p_{l'l}:\; X_{l'} \to X_l$ induces a continuous map $p_{sl}:\; Y_s \times \overline{\mathcal{G}}_{x_0}(l') \to X_l$ which satisfies

$$p_l \circ f^{-1}_{\mathcal{E}(l)} = p_{sl} \circ (p_s \times \pi_l) \tag{I.2.1. 49}$$

for an appropriate choice $\mathcal{E}(l)$ of $\mathcal{E}$.
iii)
For any two choices $\mathcal{E}, \mathcal{E}'$ the map

$$f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}:\; \overline{Y} \times \overline{\mathcal{G}}_{x_0} \to \overline{Y} \times \overline{\mathcal{G}}_{x_0} \tag{I.2.1. 50}$$

is a homeomorphism.

Proof of Lemma I.2.6:
i)
Let $s \in \mathcal{S}$ be freely generated by the independent hoops $\alpha_1, .., \alpha_m$, let $\check{\gamma}$ be the unoriented graph they determine and choose some orientation for it. Then every α_k is a finite composition of the edges $e_1, .., e_n \in E(\gamma)$ demonstrating that s is a subgroup of $l = l(\gamma)$ consisting of hoops based at $x_0 \in V(\gamma)$. We have bijections $\rho_{\alpha_1,..,\alpha_m}:\; Y_s \to G^m$ and $\rho_{e_1,..,e_n}:\; Y_s \to G^n$ as in (I.2.1. 16) which can be used to define the projection $p_{ls}:\; X_l \to Y_s$. In particular we get $X_s = Y_s$ so that p_{ls} is continuous. It follows that $p_s \circ f_{\mathcal{E}}(A) = A_s = p_{ls}(A_l) = (p_{ls} \circ p_l)(A)$ so that $p_s \circ f_{\mathcal{E}}$ is continuous.
ii)
Let $l \in \mathcal{L}$ be freely generated by independent edges $e_1, .., e_n$ and let γ be the oriented graph they determine. If $x_0 \in V(\gamma)$ invert the orientation of e_k if necessary in order to achieve that $f(e_k) \neq x_0$ for any $k = 1, .., n$. For every vertex $v \in V(\gamma)$ not yet connected to x_0 through one of the edges $e_1, .., e_n$ add another edge e_v connecting x_0 with v to the set $\{e_1, .., e_n\}$ so that the extended set remains independent. The extended set $\{e_1, .., e_{n'}\}$ determines an oriented graph γ' with $x_0 \in V(\gamma')$ and every vertex of γ' is conncted to x_0 through at least one edge. Given $v \in V(\gamma')$ choose one edge $e^l_v \in \hom(x_0, v)$ from $e_1, ., e_{n'}$ with the convention that $e^l_{x_0}$ be the trivial hoop. Define $\mathcal{E}'(l) := \{e^l_v;\; v \in V(\gamma) \cup \{x_0\}\}$ and let $\{e'_1, .., e'_m\} := \{e_1, .., e_{n'}\} - \mathcal{E}'(l)$. The hoops based at x_0 given by $\alpha_k := e^l_{b(e'_k)} \circ e'_k \circ (e^l_{f(e'_k)})^{-1},\; k = 1, .., m$ are independent due to the segments e'_k traversed precisely once and which are intersected by the other α_l in only a finite number of points (namely the end points). Let s be the subgroup of $\mathcal{Q}$ generated by the α_k and let $l' \in \mathcal{L}$ be the subgroupoid generated by the $(e^l_x)^{-1} \circ \alpha_k \circ e^l_y,\; x, y \in V(\gamma) \cup \{x_0\}, k = 1, .., m$ (we know that it is a connected subgroupoid with $\hom_{l'}(x_0, x_0) = s$ from lemma I.2.2). We claim $l \prec l'$. To see this, consider the original set of edges $\{e_1, .., e_n\}$. Each $e_k,\; k = 1, .., n$ is either one of the $e^l_v,\; v \in V(\gamma) \cup \{x_0\}$ or one of the $e'_j,\; j = 1, .., m$. In the first case we have $e_k = e^l_v = e^{-1}_{x_0} \circ e_{x_0} \circ e^l_v \in l'$ where e_{x_0} is the trivial hoop. In the latter case by definition $e_k = e'_j = (e^l_{b(e'_j)})^{-1} \circ \alpha_j \circ e^l(f(e'_j)) \in l'$.

Consider now the bijection

$$f^{l'}_{\mathcal{E}'(l)}:\; X_{l'} \to Y_s \times \overline{\mathcal{G}}_{x_0}(l') \tag{I.2.1. 51}$$

defined exactly as in (I.2.1. 43) but restricted to $X_{l'}$ so that only the system of edges $\mathcal{E}'(l)$ is needed in order to define it. We can define now

$$p_{sl} := p_{l'l} \circ (f^{l'}_{\mathcal{E}'(l)})^{-1} : \; Y_s \times \overline{\mathcal{G}}_{x_0}(l') \to X_l \tag{I.2.1. 52}$$

which is trivially continuous again because both X_l and $Y_s \times \overline{\mathcal{G}}_{x_0}(l')$ are identified with powers of G.

Let finally $\mathcal{E}(l)$ be any system of paths $e_x \in \hom(x_0, x)$ that contains $\mathcal{E}'(l)$. Then for any $B \in \mathrm{Hom}(\mathcal{Q}, G), g \in \overline{\mathcal{G}}_{x_0}, p \in l$ we have

$$\begin{aligned} & [(p_l \circ f^{-1}_{\mathcal{E}(l)})(B,g)](p) = [f^{-1}_{\mathcal{E}(l)}(B,g)](p) = g(b(p))^{-1} B(e^l_{b(p)} \circ p \circ (e^l_{f(p)})^{-1}) g(f(p)) \\ = & \; (\pi_l \circ g)(b(p))^{-1} (p_s \circ B)(e^l_{b(p)} \circ p \circ (e^l_{f(p)})^{-1}) (\pi_l \circ g)(f(p)) \\ = & \; [(f^{l'}_{\mathcal{E}'(l)})^{-1})(p_s \circ B, \pi_l \circ g)](p) = (p_{l'l} \circ (f^{l'}_{\mathcal{E}'(l)})^{-1})(p_s \circ B, \pi_l \circ g)](p) \\ = & \; [p_{sl} \circ (p_s \times \pi_l)(B,g)](p) \end{aligned} \tag{I.2.1. 53}$$

where in the second line we exploited that $b(p), f(p) \in V(\gamma)$ and that $e^l_{b(p)} \circ p \circ (e^l_{f(p)})^{-1} \in s$, in the third we observed that only the subset $\mathcal{E}'(l) \subset \mathcal{E}(l)$ is being used and that $p \in l \prec l'$ and finally we used (I.2.1. 52). Thus, $p_l \circ f^{-1}_{\mathcal{E}(l)} = p_{sl} \circ (p_s \times \pi_l)$ is a composition of continuous maps and therefore continuous.

iii)
Let $\mathcal{E} = \{e_x, \; x \in \sigma\}, \mathcal{E}' = \{e'_x, \; x \in \sigma\}$ and $\alpha \in \mathcal{Q}, x \in \sigma$, then

$$\begin{aligned} & [f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}(B,g)](\alpha, x) = ([f^{-1}_{\mathcal{E}'}(B,g)](\alpha), [f^{-1}_{\mathcal{E}'}(B,g)](e_x)) \\ = & \; (g(b(\alpha))^{-1} B(e'_{b(\alpha)} \circ \alpha \circ (e'_{f(\alpha)})^{-1}) g(f(\alpha)), (g(b(e_x))^{-1} B(e'_{b(e_x)} \circ e_x \circ (e'_{f(e_x)})^{-1}) g(f(e_x))) \\ = & \; (B(\alpha), B(e_x \circ (e'_x)^{-1}) g(x)) \end{aligned} \tag{I.2.1. 54}$$

where in the last step we noticed that $f(e_x) = x, \; b(\alpha) = f(\alpha) = b(e_x) = x_0, \; g(x_0) = 1_G$ because $g \in \overline{\mathcal{G}}_{x_0}$ and that e'_{x_0} is the trivial hoop based at x_0. It follows that the map (I.2.1. 50) is given by $(B, g) \mapsto (B', g')$ with $B' = B, g'(.) = B(e_x \circ (e'_x)^{-1}) g(.)$. The inverse map is given similarly by $(B, g) \mapsto (B', g')$ with $B' = B, g'(.) = B(e'_x \circ (e_x)^{-1}) g(.)$ so that it will be sufficient to demonstrate continuity of the former.

To show that $f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}$ is continuous requires to show that $(p_s \times p_x) \circ f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}$ is continuous for all $s \in \mathcal{S}, x \in \sigma$. Now obviously $p_s \circ f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'} = p_s$ is continuous by definition. Next $[p_x \circ f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}(B,g)](x) = B(e_x \circ (e'_x)^{-1}) g(x)$. Define the restriction map

$$f^{\mathcal{E},\mathcal{E}'}_x := p_{e_x \circ (e'_x)^{-1}} \times p_x : \; \overline{Y} \times \overline{\mathcal{G}}_{x_0} \to Y_{e_x \circ (e'_x)^{-1}} \times (\overline{\mathcal{G}}_{x_0})_{|x} \tag{I.2.1. 55}$$

and denote by $m : \; G \times G \to G; \; (g_1, g_2) \to g_1 g_2$ multiplication in G. Then

$$p_x \circ f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'} = m \circ (p_{e_x \circ (e'_x)^{-1}} \times p_x) \tag{I.2.1. 56}$$

is a composition of continuous maps and therefore continuous. Hence, $f_{\mathcal{E}} \circ f^{-1}_{\mathcal{E}'}$ is a homeomorphism. $\Box$

We can now complete the proof of continuity of both $f_{\mathcal{E}}$ and $f^{-1}_{\mathcal{E}}$ for a given, fixed $\mathcal{E}$. We showed already that $p_x \circ f_{\mathcal{E}}$ is continuous for all $x \in \sigma$ and by lemma I.2.6i) we have that $p_s \circ f_{\mathcal{E}}$ is continuous for all $s \in \mathcal{S}$, hence $f_{\mathcal{E}}$ is continuous. Next

$$p_l \circ f^{-1}_{\mathcal{E}} = [p_l \circ f^{-1}_{\mathcal{E}(l)}] \circ [f_{\mathcal{E}(l)} \circ f^{-1}_{\mathcal{E}}] \tag{I.2.1. 57}$$

is a composition of two continuous functions since the function in the first bracket is continuous by lemma I.2.6ii) and the second by lemma I.2.6iii), thus $f_{\mathcal{E}}^{-1}$ is continuous.
□

Theorem I.2.7
The spaces $\overline{\mathcal{A}/\mathcal{G}} = Hom(\mathcal{P}, G)/\overline{\mathcal{G}}$ and $\overline{\mathcal{A}/\mathcal{G}} = Hom(\mathcal{Q}, G)/G$ are homeomorphic.

Proof of Theorem I.2.7:
By theorem I.2.6 we know that
1) $\mathrm{Hom}(\mathcal{P}, G)$ and $\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0}$ are homeomorphic and
2) $\overline{\mathcal{G}}$ acts equivariantly on both spaces via λ, λ' respectively.
We now use the abstract result that if a group acts (not necessarily continuously) equivariantly on two homeomorphic spaces then the corresponding spaces continue to be homeomorphic in their respective quotient topologies (see section III.3). We therefore know that $\mathrm{Hom}(\mathcal{P}, G)/\overline{\mathcal{G}}$ and $(\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0})/\overline{\mathcal{G}}$ are homeomorphic. But $\overline{\mathcal{G}}$ is a direct product space, that is, $\overline{\mathcal{G}} = \overline{\mathcal{G}}_{x_0} \times G$ whence $(\mathrm{Hom}(\mathcal{Q}, G) \times \overline{\mathcal{G}}_{x_0})/\overline{\mathcal{G}} = \mathrm{Hom}(\mathcal{Q}, G)/G$. More explicitly, recalling the action of λ' in (I.1.1. 41) and writing $g \in \overline{\mathcal{G}}$ as $g = (g_1, g_0) \in \overline{\mathcal{G}}_{x_0} \times G$ where $g(x) = g_1(x)$ for $x \neq x_0$ and $g(x_0) = g_0$ we see that $B^g(\alpha) = \mathrm{Ad}_{g_0}(B(\alpha))$ and $h^g(x) = g_0 h(x) g(x)^{-1}$ which gives $h^g(x_0) = h(x_0) = 1_G$ and $h^g(x) = g_0 h(x) g_1(x)^{-1}$ for $x \neq x_0$. It follows that, given $h \in \overline{\mathcal{G}}_{x_0}$, for any choice of g_0 we can gauge $h^g(x) = 1_G$ for all $x \in \sigma$ by choosing $g_1(x) = g_0 h(x)$. The remaining gauge freedom expressed in g_0 then only acts by Ad on $\mathrm{Hom}(\mathcal{Q}, G)$.
□

I.2.2 The C^* Algebraic Viewpoint

In the previous sections we have defined the quantum configuration spaces of (gauge equivalence casses of) distributional connections $\overline{\mathcal{A}}$ ($\overline{\mathcal{A}/\mathcal{G}}$) as $\mathrm{Hom}(\mathcal{P}, G)$ ($\mathrm{Hom}(\mathcal{P}, G)/\overline{\mathcal{G}}$) and equipped them with the Tychonov topology through projective techniques. We could be satisfied with this because we know that these spaces are compact Hausdorff spaces and this is a sufficiently powerful result in order to develop measure theory on them as we will se below.

However, the result that we want to establish in this section, namely that both spaces can be seen as the Gel'fand spectra of certain C^* algebras, has the advantage to make the connection with so-called cylindrical functions on these spaces explicit which then helps to construct (a priori only cylindrically defined) measures on them. Moreover, it has a wider range of applicability in the sense that it does not make use of the concrete label sets used in the previous section. It therefore establishes a concrete link with constructive quantum gauge field theories. A brief introduction to Gel'fand – Naimark – Segal theory can be found in section III.4. We will follow closely Ashtekar and Lewandowski [109]

We begin again quite generally and suppose that we are given a partially ordered and directed index set $\mathcal{L}$ which label compact Hausdorff spaces X_l and that we have surjective and continuous projections $p_{l'l} : X_{l'} \to X_l$ for $l \prec l'$ satisfying the consistency condition $p_{l'l} \circ p_{l^dprimel'} = p_{l''l}$ for $l \prec l' \prec l''$. Let $X_\infty, \overline{X}$ be the corresponding direct product and projective limit respectively with Tychonov topology with respect to which we know that they are Hausdorff and compact from the previous sections.

Definition I.2.16
i)

Let $C(X_l)$ be the continuous, complex valued functions on X_l and consider their union

$$\mathit{Cyl}'(\overline{X}) := \cup_{l\in\mathcal{L}} C(X_l) \tag{I.2.2. 1}$$

Given $f, f' \in \mathit{Cyl}(\overline{X})$ we find $l, l' \in \mathcal{L}$ such that $f \in C(X_l), f' \in C(X_{l'})$ and we say that f, f' are equivalent, denoted $f \sim f'$ provided that

$$p^*_{l^d primel} f = p_{l''l'} f' \; \forall \; l, l' \prec l'' \tag{I.2.2. 2}$$

(pull-back maps) ii)
The space of cylindrical functions on the projective limit $\overline{X}$ is defined to be the space of equivalence classes

$$\mathit{Cyl}(\overline{X}) := \mathit{Cyl}'(\overline{X})/\sim \tag{I.2.2. 3}$$

We will denote the equivalence class of $f \in \mathit{Cyl}'(\overline{X})$ by $[f]_\sim$.

Notice that we are actually abusing the notation here since an element $f \in \mathrm{Cyl}(\overline{X})$ is not a function on $\overline{X}$ but an equivalence class of functions on the X_l. We will justify this later by showing that $\mathrm{Cyl}(\overline{X})$ can be identified with $C(\overline{X})$, the continuous functions on $\overline{X}$.

Condition (I.2.2. 2) seems to be very hard to check but it is sufficient to find just one single l'' such that (I.2.2. 2). For suppose that $f_{l_1} \in C(X_{l_1}), f_{l_2} \in C(X_{l_2})$ are given and that we find some $l_1, l_2 \prec l_3$ such that $p^*_{l_3 l_1} f_{l_1} = p^*_{l_3 l_2} f_{l_2}$ Now let any $l_1, l_2 \prec l_4$ be given. Since $\mathcal{L}$ is directed we find $l_1, l_2, l_3, l_4 \prec l_5$ and due to the consistency condition among the projections we have

$$\text{i) } p_{l_4 l_1} \circ p_{l_5 l_4} = p_{l_5 l_1} = p_{l_3 l_1} \circ p_{l_5 l_3} \text{ and ii) } p_{l_4 l_2} \circ p_{l_5 l_4} = p_{l_5 l_2} = p_{l_3 l_2} \circ p_{l_5 l_3} \tag{I.2.2. 4}$$

whence

$$p^*_{l_5 l_4} p^*_{l_4 l_1} f_{l_1} =_{\mathrm{i)}} \; p^*_{l_5 l_3} p^*_{l_3 l_1} f_{l_1} = p^*_{l_5 l_3} p^*_{l_3 l_2} f_{l_2} =_{\mathrm{ii)}} \; p^*_{l_5 l_4} p^*_{l_4 l_2} f_{l_2} \tag{I.2.2. 5}$$

where in the middle equality we have used (I.2.2. 2) for $l'' = l_3$. We conclude that $p^*_{l_5 l_4}[p^*_{l_4 l_1} f_{l_1} - p^*_{l_4 l_2} f_{l_2}] = 0$. Now for any $f_{l_4} \in C(X_{l_4})$ the condition $f_{l_4}(p_{l_5 l_4}(x_{l_5})) = 0$ for all $x_{l_5} \in X_{l_5}$ means that $f_{l_4} = 0$ because $p_{l_5 l_4} : \; X_{l_5} \to X_{l_4}$ is surjective.

Lemma I.2.7
Given $f, f' \in \mathit{Cyl}(\overline{X})$ there exists a common label $l \in \mathcal{L}$ and $f_l, f'_l \in C(X_l)$ such that $f = [f_l]_\sim, f' = [f'_l]_\sim$.

Proof of Lemma I.2.7:
By definition we find $l_1, l_2 \in \mathcal{L}$ and representatives $f_{l1} \in C(X_{l1}), f_{l2} \in C(X_{l2})$ such that $f = [f_{l_1}]_\sim, f' = [f_{l_2}]_\sim$. Choose any $l_1, l_2 \prec l$ then $f_l := p^*_{l l_1} f_{l1} \sim f_{l_1}$ (choose $l'' = l$ in (I.2.2. 2 and use $p_{ll} = \mathrm{id}_{X_l}$) and $f'_l := p^*_{l l_2} f_{l2} \sim f_{l_2}$. Thus $f = [f_l]_\sim, f' = [f'_l]_\sim$.
$\square$

Lemma I.2.8 *I.2.9*
i)
Let $f, f' \in \mathit{Cyl}(\overline{X})$ then the following operations are well defined (independent of the representatives)

$$f + f' := [f_l + f'_l]_\sim, \; ff' := [f_l f'_l]_\sim, \; zf := [zf_l]_\sim, \; \bar{f} := [\bar{f}_l]_\sim \tag{I.2.2. 6}$$

where l, f_l, f'_l are as in lemma I.2.7, $z \in \mathbb{C}$ and $\bar{f}_l$ denotes complex conjugation.
ii)

Cyl($\overline{X}$) contains the constant functions.
iii)
The sup − norm for $f = [f_l]_\sim$

$$||f|| := \sup_{x_l \in X_l} |f_l(x_l)| \tag{I.2.2. 7}$$

is well-defined.

Proof of Lemma I.2.9:
i)
We consider only pointwise multiplication, the other cases are similar. Let l, f_l, f'_l and $l', f_{l'}, f'_{l'}$ as in lemma I.2.7. We find $l, l' \prec l''$ and have $p^*_{l''l} f_l = p^*_{l''l'} f_{l'}$ and $p^*_{l''l} f'_l = p^*_{l''l'} f'_{l'}$. Thus

$$p^*_{l''l}(f_l f'_l) = p^*_{l''l}(f_l) p^*_{l''l}(f'_l) = p^*_{l''l'}(f_{l'}) p^*_{l''l'}(f'_{l'}) = p^*_{l''l'}(f_{l'} f'_{l'}) \tag{I.2.2. 8}$$

so $f_l f'_l \sim f_{l'} f'_{l'}$.
ii)
The function $f^z_l : X_l \to \mathbb{C};\ x_l \to z$ for any $z \in \mathbb{C}$ certainly is an element of $C(X_l)$ and for any $l, l' \prec l''$ we have $z = (p^*_{l''l} f^z_l)(x_{l''}) = (p^*_{l''l'} f^z_{l'})(x_{l''})$ for all $x_{l''} \in X_{l''}$ so $f^z := [f^z_l]_\sim$ is well-defined.
iii)
If $f = [f_l]_\sim = [f_{l'}]_\sim$ is given, choose any $l, l' \prec l''$ so that we know that $p^*_{l''l} f_l = p^*_{l''l'} f_{l'}$. Then from the surjectivity of $p_{l''l}, p_{l''l'}$ we have

$$\sup_{x_l \in X_l} |f_l(x_l)| = \sup_{x_{l''} \in X_{l''}} |(p^*_{l''l} f_l)(x_{l''})| = \sup_{x_{l''} \in X_{l''}} |(p^*_{l''l'} f_{l'})(x_{l''})| = \sup_{x_{l'} \in X_{l'}} |f_{l'}(x_{l'})| \tag{I.2.2. 9}$$

□

Lemma I.2.9i) tells us that Cyl($\overline{X}$) is an Abelean, $^*-$algebra defined by the pointwise operations I.2.2. 6. Lemma I.2.9ii) tells us that Cyl($\overline{X}$) is also unital, the unit being given by the constant function $1 = [1_l]_\sim$, $1_l(x_l) = 1$. Finally, lemma I.2.9iii) tells us that Cyl($\overline{X}$) is a normed space and that the norm is correctly normalized, that is, $||1|| = 1$. Notice that here the compactness of the X_l comes in since the norm (I.2.2. 7) certainly does not make sense any longer on $C(X_l)$ for non-compact X_l. If X_l is at least locally compact we can replace the $C(X_l)$ by $C_0(X_l)$, the continuous complex valued functions of compact support and still would get an Abelean * algebra with norm although no longer a unital one. One can always embed an algebra isometrically into a larger algebra with identity (even preserving the C^* property, see below) but this does not solve all problems in C^*-algebra theory. Fortunately, we have not to deal with these complications in what follows.

Recall that a norm induces a metric on a linear space via $d(f, f') := ||f - f'||$ and that a metric space is said to be complete whenever all its Cauchy sequences converge. Any incomplete metric space can be uniquely (up to isometry) embedded into a complete metric space by extending it by its non-converging Cauchy sequences (see e.g. [129]). We can then complete Cyl($\overline{X}$) in the norm $||.||$ in this sense and obtain an Abelean, unital Banach $^*-$algebra $\overline{\text{Cyl}(\overline{X})}$. But we notice that not only the submultiplicativity of the norm ($||ff'|| \le ||f||\ ||f'||$) holds but in fact the C^* property $||f\bar{f}|| = ||f||^2$. Thus $\overline{\text{Cyl}(\overline{X})}$ is in fact an unital, Abelean C^*-algebra. This observation suggests to apply Gel'fand-Naimark-Segal theory to which an elementary introduction can be found in section III.4.

Denote by $\Delta(\text{Cyl}(\overline{X}))$ the spectrum of $\text{Cyl}(\overline{X})$, that is, the set of *all* (algebraic, i.e. not necessarily continuous) homomorphism from $\text{Cyl}(\overline{X})$ into the complex numbers and denote the Gel'fand isometric isomorphism by

$$\bigvee :\ \overline{\text{Cyl}(\overline{X})} \to C(\Delta(\overline{\text{Cyl}(\overline{X})}));\ f \mapsto \check{f} \text{ where } \check{f}(\chi) := \chi(f) \tag{I.2.2. 10}$$

where the space of continuous functions on the spectrum is equipped with the sup-norm. The spectrum is automatically a compact Hausdorff space in the Gel'fand topology, the weakest topology in which all the $\check{f}$, $f \in \overline{\mathrm{Cyl}(\overline{X})}$ are continuous.

Notice the similarity between the spaces $\overline{\mathrm{Cyl}(\overline{X})}$ and $C(\Delta(\overline{\mathrm{Cyl}(\overline{X})}))$: both are spaces of continuous functions over compact Hausdorff spaces and on both spaces the norm is the sup-norm. This suggests that there is a homeomorphism between the projective limit space $\overline{X}$ and the spectrum $\mathrm{Hom}(\overline{\mathrm{Cyl}(\overline{X})}, \mathbb{C})$. This is what we are going to prove in what follows.

Consider the map

$$\mathcal{X}:\ \overline{X} \to \Delta(\overline{\mathrm{Cyl}(\overline{X})});\ x = (x_l)_{l\in\mathcal{L}} \mapsto \mathcal{X}(x) \text{ where } [\mathcal{X}(x)](f) := f_l(p_l(x)) \text{ for } f = [f_l]_\sim \quad \text{(I.2.2. 11)}$$

Notice that (I.2.2. 11) is well-defined since $f = p_l^* f_l = p_{l'}^* f_{l'}$ for any $f_l \sim f_{l'}$ which follows from

$$p_l^* f_l(x) = f_l(x_l) = (p_{l''l}^* f_l)(x_{l''}) = (p_{l''l'}^* f_{l'})(x_{l''}) = f_{l'}(x_{l'}) = p_{l'}^* f_{l'}(x) \quad \text{(I.2.2. 12)}$$

for any $x \in \overline{X},\ l, l' \prec l''$. Notice also that (I.2.2. 11) a priori defines $\mathcal{X}(x)$ only on $\mathrm{Cyl}(\overline{X})$ and not on the completion $\overline{\mathrm{Cyl}(\overline{X})}$. We now show that every $\mathcal{X}(x)$ is actually continuous: Let (f^α) be a net converging in $\mathrm{Cyl}(\overline{X})$ to f, that is, $\lim_\alpha ||f_\alpha - f|| = 0$. Then $(f^\alpha = [f^\alpha_{l_\alpha}]_\sim,\ f = [f_l]_\sim,\ l, l_\alpha \prec l_{\alpha,l})$

$$\begin{aligned} |[\mathcal{X}(x)](f^\alpha) - [\mathcal{X}(x)](f)| &= |(p_{l_\alpha}^* f^\alpha_{l_\alpha} - p_l^* f_l)(x)| = |(p_{l_{\alpha,l} l_\alpha}^* f^\alpha_{l_\alpha} - p_{l_{\alpha,l}}^* f_l)(x_{l_{\alpha,l}})| \\ &= |f^\alpha_{l_{\alpha,l}} - f_{l_{\alpha,l}})(x_{l_{\alpha,l}})| \le \sup_{x_{l_{\alpha,l}} \in X_{l_{\alpha,l}}} |(f^\alpha_{l_{\alpha,l}} - f_{l_{\alpha,l}})(x_{l_{\alpha,li}})| = ||f^\alpha - f|| \end{aligned} \quad \text{(I.2.2. 13)}$$

hence $\lim_\alpha [\mathcal{X}(x)](f^\alpha) = [\mathcal{X}(x)](f)$ so $\mathcal{X}(x)$ is continuous. It follows that $\mathcal{X}(x)$ is a continuous linear (and therefore bounded) map from the normed linear space $\mathrm{Cyl}(\overline{X})$ to the complete, normed linear space $\mathbb{C}$. Hence, by the bounded linear transformation theorem [129] each $\mathcal{X}(x)$ can be uniquely extended to a bounded linear transformation (with the same bound) from the completion $\overline{\mathrm{Cyl}(\overline{X})}$ of $\mathrm{Cyl}(\overline{X})$ to $\mathbb{C}$ by taking the limit of the evaluation on convergent series in $\overline{\mathrm{Cyl}(\overline{X})}$ which are only Cauchy in $\mathrm{Cyl}(\overline{X})$. We will denote the extension of $\mathcal{X}(x)$ to $\overline{\mathrm{Cyl}(\overline{X})}$ by $\mathcal{X}(x)$ again and it is then easy to check that this extended map $\mathcal{X}$ is an element of $\Delta(\overline{\mathrm{Cyl}(\overline{X})})$ (a homomorphism), e.g. if $f_n \to f, f'_n \to f'$ then

$$[\mathcal{X}(x)](ff') := \lim_{n\to\infty} [\mathcal{X}(x)](f_n f'_n) = \lim_{n\to\infty} ([\mathcal{X}(x)](f_n))\,([\mathcal{X}(x)](f'_n)) = ([\mathcal{X}(x)](f))\,([\mathcal{X}(x)](f')) \quad \text{(I.2.2. 14)}$$

The map $\mathcal{X}$ in (I.2.2. 11) is to be understood in this extended sense.

Theorem I.2.8
The map $\mathcal{X}$ in I.2.2. 11 is a homeomorphism.

Proof of Theorem I.2.8:
Injectivity:
Suppose $\mathcal{X}(x) = \mathcal{X}(x')$, then in particular $[\mathcal{X}(x)](f) = [\mathcal{X}(x')](f)$ for any $f \in \mathrm{Cyl}(\overline{X})$. Hence $f_l(x_l) = f_l(x'_l)$ for any $f_l \in C(X_l),\ l \in \mathcal{L}$. Since X_l is a compact Hausdorff space, $C(X_l)$ separates the points of X_l by the Stone-Weierstrass theorem [129], hence $x_l = x'_l$ for all $l \in \mathcal{L}$. It follows that $x = x'$.
Surjectivity:
Let $\chi \in \mathrm{Hom}(\overline{\mathrm{Cyl}(\overline{X})}, \mathbb{C})$ be given. We must construct $x^\chi \in \overline{X}$ such that $\mathcal{X}(x^\chi) = \chi$. In particular for any $f = [f_l]_\sim \in \mathrm{Cyl}(\overline{X})$ we have $f_l(x^\chi_l) = \chi([f_l]_\sim)$. Given $l \in \mathcal{L}$ the character χ defines an element

$\chi_l \in \mathrm{Hom}(C(X_l), \mathbb{C})$ via $\chi_l(f_l) := \chi([f_l]_\sim)$ for all $f_l \in C(X_l)$. Since X_l is a compact Hausdorff space, it is the spectrum of the Abelean, unital C^*-algebra $C(X_l)$, hence $X_l = \mathrm{Hom}(C(X_l), \mathbb{C})$ (see section III.4). It follows that there exists $x_l^\chi \in X_l$ such that $\chi_l(f_l) = f_l(x_l^\chi)$ for all $f_l \in C(X_l)$. We define $x^\chi := (x_l^\chi)_{l\in\mathcal{L}}$ and must check that it defines an element of the projective limit.

Let $l \prec l'$ and $f = [f_l]_\sim$. Then $f_l \sim f_{l'} := p_{l'l}^* f_l$ (choose $l'' = l'$ and use $p_{l'l'} = \mathrm{id}_{X_{l'}}$) and therefore

$$f_l(x_l^\chi) = \chi_l(f_l) = \chi([f_l]_\sim) = \chi([f_{l'}]_\sim) = \chi_{l'}(f_{l'}) = f_{l'}(x_{l'}^\chi) = f_l(p_{l'l}(x_{l'}^\chi)) \qquad \text{(I.2.2. 15)}$$

for any $f_l \in C(X_l)$, $l \in \mathcal{L}$. Since $C(X_l)$ separates the points of X_l we conclude $x_l^\chi = p_{l'l}(x_{l'}^\chi)$ for any $l \prec l'$, hence $x^\chi \in \overline{X}$.

Continuity:

We have established that $\mathcal{X}$ is a bijection. We must show that both $\mathcal{X}, \mathcal{X}^{-1}$ are continuous.

The topology on $\Delta(\overline{\mathrm{Cyl}(\overline{X})})$ is the weakest topology such that the Gel'fand transforms $\check{f},\ f \in \overline{\mathrm{Cyl}(\overline{X})})$ are continuous while the topology on $\overline{X}$ is the weakest topology such that all the projections p_l are continuous, or equivalently that al the $p_l^* f_l,\ f_l \in C(X_l)$ are continuous.

Continuity of $\mathcal{X}$:

Let (x^α) be a net in $\overline{X}$ converging to x, that is, every net (x_l^α) converges to x_l. Let first $f = [f_l]_\sim \in \mathrm{Cyl}(\overline{X})$. Then

$$\lim_\alpha [\mathcal{X}(x^\alpha)](f) = \lim_\alpha (p_l^* f_l)(x^\alpha) = (p_l^* f_l)(x) = [\mathcal{X}(x)](f) \qquad \text{(I.2.2. 16)}$$

for any $f \in \mathrm{Cyl}(\overline{X})$. Now given $\epsilon > 0$ for general $f \in \overline{\mathrm{Cyl}(\overline{X})}$ we find $f_\epsilon \in \mathrm{Cyl}(\overline{X})$ such that $||f - f_\epsilon|| < \epsilon/3$ because $\mathrm{Cyl}(\overline{X})$ is dense in $\overline{\mathrm{Cyl}(\overline{X})}$. Also, by (I.2.2. 16), we find $\alpha(\epsilon)$ such that $|[\mathcal{X}(x^\alpha)(f_\epsilon) - [\mathcal{X}(x)](f_\epsilon)| \le \epsilon/3$ for any $\alpha(\epsilon) \prec \alpha$. Finally, since $\mathcal{X}(x^\alpha), \mathcal{X}(x)$ are characters they are bounded (by one) linear functionals on $\overline{\mathrm{Cyl}(\overline{X})}$ as we have shown above (continuity of the $\mathcal{X}(x)$). It follows that

$$\begin{aligned} |[\mathcal{X}(x^\alpha)](f) - [\mathcal{X}(x)](f)| &\le |[\mathcal{X}(x^\alpha)](f - f_\epsilon)| + |[\mathcal{X}(x)](f - f_\epsilon)| + |[\mathcal{X}(x^\alpha)](f_\epsilon) - [\mathcal{X}(x)](f_\epsilon)| \\ &\le 2||f - f_\epsilon|| + \epsilon/3 \le \epsilon \end{aligned} \qquad \text{(I.2.2. 17)}$$

for all $\alpha(\epsilon) \prec \alpha$. Thus

$$\lim_\alpha \check{f}(\mathcal{X}(x^\alpha)) = \check{f}(\mathcal{X}(f)) \qquad \text{(I.2.2. 18)}$$

for all $f \in \overline{\mathrm{Cyl}(\overline{X})}$, hence $\mathcal{X}(x^\alpha) \to \mathcal{X}(x)$ in the Gel'fand topology.

$\mathcal{X}^{-1}$:

Let (χ^α) be a net in $\Delta(\overline{\mathrm{Cyl}(\overline{X})})$ converging to χ, so $\chi^\alpha(f) \to \chi(f)$ for any $f \in \overline{\mathrm{Cyl}(\overline{X})}$ and so in particular for $f = [f_l]_\sim \in \mathrm{Cyl}(\overline{X})$. Therefore

$$\chi^\alpha(f) = \chi^\alpha(p_l^* f_l) = (p_l^* f_l)(x^{\chi_\alpha}) = (p_l^* f_l)(\mathcal{X}^{-1}(\chi_\alpha)) \to (p_l^* f_l)(\mathcal{X}^{-1}(\chi)) = \chi(f) \qquad \text{(I.2.2. 19)}$$

for all $f_l \in C(X_l), l \in \mathcal{L}$. Hence $\mathcal{X}^{-1}(\chi_\alpha) \to \mathcal{X}^{-1}(\chi)$ in the Tychonov topology.

□

Corollary I.2.1

The closure of the space of cylindrical functions $\overline{Cyl(\overline{X})}$ may be identified with the space of continuous functions $C(\overline{X})$ on the sprojective limit $\overline{X}$.

This follows from the fact that via theorem I.2.8 we may identify $\overline{X}$ set-theoretically and topologically with the spectrum $\Delta(\overline{\mathrm{Cyl}(\overline{X})})$ and the fact that the Gel'fand transform between $\overline{\mathrm{Cyl}(\overline{X})}$ and

$C(\Delta(\overline{\mathrm{Cyl}(\overline{X})}))$ is an (isometric) isomorphism. This justifies in retrospect the notation $\mathrm{Cyl}(\overline{X})$ although cylindrical functions are not functions on $\overline{X}$ but rather equivalence classes of functions on the X_l under $\sim$.

Next we give an abstract and independent C^*-algebraic proof for the fact that the spaces $\overline{X}/G$ and $\overline{X/G}$ are homeomorphic whenever a topological group G acts continuously and equivariantly on the projective limit $\overline{X}$, that is, we reprove theorem I.2.7.

Suppose then that for each $l \in \mathcal{L}$ we have a group action

$$\lambda^l : \; G \times X_l \to X_l; \; (g, x_l) \mapsto \lambda^l_g(x_l) \tag{I.2.2. 20}$$

where λ^l_g is a continuous map on X_l which is equivariant with repspect to the projective structure, that is,

$$p_{l'l} \circ \lambda^{l'} = \lambda^l \circ p_{l'l} \; \forall l \prec l' \tag{I.2.2. 21}$$

Due to continuity of the group action and since X_l is Hausdorff and compact, the quotient space X_l/G is again compact and Hausdorff in the quotient topology (see section III.3) and due to equivariance the net of equivalence classes $([x_l]_l)_{l\in\mathcal{L}}$ is a projective net again (with respect to the same projections $p_{l'l}$) so that we can form the projective limit $\overline{X/G}$ of the X_l/G which then is a compact Hausdorff space again. Here $[.]_l : \; X_l \to X_l/G$ denotes the individual quotient maps with respect to the λ^l.

On the other hand we may directly define an action of G on $\overline{X}$ itself by

$$\lambda : \; \overline{X} \times G \to \overline{X}; \; x = (x_l)_{l\in\mathcal{L}} \mapsto \lambda_g(x) := (\lambda^l_g(x_l))_{l\in\mathcal{L}} \tag{I.2.2. 22}$$

Since $\overline{X}$ is compact and Hausdorff and λ_g is a continuous map on $\overline{X}$ (since it is continuous iff all the λ^l_g are continuous) it follows that the quotient space $\overline{X}/G$ is again a compact Hausdorff space.

We now want to know what the relation between $\overline{X}/G$ and $\overline{X/G}$ is. Let $[.] : \; \overline{X} \to \overline{X}/G$ be the quotient map with respect to λ. We then may define a map

$$\Phi : \; \overline{X}/G \to \overline{X/G}; \; [x] = [(x_l)_{l\in\mathcal{L}}] \mapsto ([x_l]_l)_{l\in\mathcal{L}} \tag{I.2.2. 23}$$

as follows: we have

$$[x] = \{\lambda_g(x); \; g \in G\} := \{(\lambda^l_g(x_l))_{l\in\mathcal{L}} \; g \in G\} \tag{I.2.2. 24}$$

Now take an arbitrary representative in $[x]$, say $\lambda_{g_0}(x)$ for some $g_0 \in G$ and compute its class in $\overline{X/G}$, that is,

$$\Phi([x]) := ([p_l(\lambda_{g_0}(x))]_l)_{l\in\mathcal{L}} = (\{\lambda^l_g(\lambda^l_{g_0}(x_l)); \; g \in G\})_{l\in\mathcal{L}} = (\{\lambda^l_g(x_l); \; g \in G\})_{l\in\mathcal{L}} \tag{I.2.2. 25}$$

which shows that Φ is well-defined, that is, independent of the choice of g_0.

Theorem I.2.9
The map Φ defined in (I.2.2. 23) is a homeomorphism.

Proof of Theorem I.2.9:
The strategy of the proof is to 1) first show that the pull-back map

$$\Phi^* : \; C(\overline{X/G}) \to C(\overline{X}/G) \tag{I.2.2. 26}$$

is a bijection and then 2) to show that for any compact Hausdorff spaces A, B such that $\Phi^* : \; A \to B$ is a bijection it follows that Φ is a homeomorphism.

Step 1)
Let $f \in C(\overline{X/G})$ be given. Via corollary I.2.1 we may think of f as an element of $\overline{\text{Cyl}(\overline{X/G})}$ and elements of $\text{Cyl}(\overline{X/G})$ lie dense in that space. Now any $f \in \text{Cyl}(\overline{X/G})$ is given by $f = [f_l]_\sim$ where f_l is a λ^l invariant function on X_l. Then

$$f_l([x_l]_l) = p_l^* f_l(\Phi([x])) \qquad \text{(I.2.2. 27)}$$

Thus the functions on $\text{Cyl}(\overline{X/G})$ are obtained as $p_l^* f_l$ for some $l \in \mathcal{L}$ where f_l is λ^l invariant and then $\Phi^* p_l^* f_l$ is a $\lambda-$invariant function on $\overline{X}$. But such functions are precisely those that lie dense in $C(\overline{X}/G)$ because a function $f \in C(\overline{X}/G)$ is simply a $\lambda-$invariant function in $C(\overline{X})$, that is, via corollary I.2.1 a $\lambda-$invariant function in $\overline{\text{Cyl}(\overline{X})}$ in which the $\lambda-$invariant functions in $\text{Cyl}(\overline{X})$ lie dense and the latter are of the form $p_l^* f_l$ for some $l \in \mathcal{L}$ and $\lambda-$invariant.

To see that Φ^* is injective on $\text{Cyl}(\overline{X/G})$ suppose that $\Phi^* p_l^* f_l = \Phi^* p_{l'}^* f'_{l'}$ for some l, l'. Then trivially $p_l^* f_l(x) = p_{l'}^* f'_{l'}(x)$ for all $x \in \overline{X}$. Let $l, l' \prec l''$ then

$$p_l^* f_l(x) = f_l(x_l) = p_{l''l}^* f_l(x_{l''}) = p_{l'}^* f'_{l'}(x) = f_{l'}(x_{l'}) = p_{l''l'}^* f_{l'}(x_{l''}) \;\forall\; x_{l''} \in X_{l''} \qquad \text{(I.2.2. 28)}$$

which shows that $f_l \sim f'_{l'}$, hence $[f_l]_\sim = [f'_{l'}]_\sim$ define the same element of $\text{Cyl}(\overline{X/G})$.

To see that Φ^* is a surjection we notice that it maps the dense set of functions in $\overline{\text{Cyl}(\overline{X/G})}$ of the form $p_l^* f_l$ (f_l being λ^l-invariant) into the dense set of functions in $\overline{\text{Cyl}(\overline{X}/G)}$ of the form $\Phi^* p_l^* f_l$ that are $\lambda-$invariant. If we can show that $\Phi^* :\; \text{Cyl}(\overline{X/G}) \to \overline{\text{Cyl}(\overline{X}/G)}$ is continuous then it can be uniquely extended as a continuous map to the completion $\Phi^* :\; \overline{\text{Cyl}(\overline{X/G})} \to \overline{\text{Cyl}(\overline{X}/G)}$ by the bounded linear transformation theorem and it will be a surjection since any $f \in \overline{\text{Cyl}(\overline{X}/G)}$ can be approximated arbitrarily well by elements in $\text{Cyl}(\overline{X}/G)$ which we know to lie in the image of Φ^* already. To prove that Φ^* is continuous (bounded), we show that it is actually an isometry and therefore has unity bound.

$$\begin{aligned} ||\Phi^* p_l^* f_l||_{\text{Cyl}(\overline{X}/G)} &= \sup_{[x]\in\overline{X}/G} |f_l(p_l(\Phi([x])))| \\ &= \sup_{([x_{l'}]_{l'})_{l'\in\mathcal{L}}\in\overline{X/G}} |f_l(p_l(([(x_{l'}]_{l'})_{l'\in\mathcal{L}})))| = ||p_l^* f_l||_{\overline{\text{Cyl}(\overline{X/G})}} \end{aligned} \qquad \text{(I.2.2. 29)}$$

Step 2)
Let $\Phi :\; A \to B$ be a map between compact Hausdorff spaces such that $\Phi^* :\; C(B) \to C(A)$ is a bijection.
Injectivity:
Suppose $\Phi(a) = \Phi(a')$. Then for any $F \in C(B)$ we have $(\Phi^* F)(a) = (\Phi^* F)(a')$. Since Φ^* is a surjection and $C(A)$ separates the points of A it follows that $a = a'$.
Surjectivity:
Since A, B are the Gel'fand spectra $\text{Hom}(C(A), \mathbb{C}), \text{Hom}(C(B), \mathbb{C})$ of $C(A), C(B)$ respectively and Φ^* is a bijection we obtain a corresponding bijection between A, B (since the spectrum can be constructed algebraically from the algebras) via

$$\Phi_* :\; A = \Delta(C(A)) \to B = \Delta(C(B)); a \mapsto a \circ \Phi^* \qquad \text{(I.2.2. 30)}$$

where

$$f(a) \equiv a(f) = a(\Phi^* F) = (a \circ \Phi^*)(F) = F(\Phi(a)) = (\Phi(a))(F) \qquad \text{(I.2.2. 31)}$$

for any $f = \Phi^* F \in C(A),\; F \in C(B)$. It follows that any $b \in B$ can be written in the form $b = \Phi(a)$ for some $a \in A$.

Continuity:
We know that both $\Phi^{-1}, (\Phi^*)^{-1}$ exist. Then $(\Phi^*)^{-1} = (\Phi^{-1})^*$ since

$$\begin{aligned} f(a) &= [(\Phi^* \circ (\Phi^*)^{-1})f](a) = [(\Phi^*)^{-1}f](\Phi(a)) \\ &= f((\Phi^{-1} \circ \Phi)(a)) = [(\Phi^{-1})^* f](\Phi(a)) \end{aligned} \tag{I.2.2. 32}$$

for any $f \in C(A), a \in A$. Let now (a^α) be a net in A converging to a. This is equivalent with $\lim_\alpha f(a^\alpha) = f(a)$ for all $f \in C(A)$ which in turn implies $\lim_\alpha F(\Phi(a^\alpha)) = F(\Phi(a))$ for all $F \in C(B)$ since any f can be written as $\Phi^* F$ which then is equvalent with the convergence of the net $\Phi(a^\alpha)$ to $\Phi(a)$ in B. The proof for Φ^{-1} is anlogous.
□

I.2.3 Regular Borel Measures on the Projective Limit: The Uniform Measure

In this section we describe a simple mechanism, based on the Riesz representation theorem, of how to construct $\sigma-$additive measures on the projective limit $\overline{X}$ starting from a so-called self-consistent family of (so-called cylindrical) measures μ_l on the various X_l. See section III.5 for some useful measure theoretic terminology and the references cited there for further reading.

Our spaces X_l are compact Hausdorff spaces and in particular topological spaces and are therefore naturally equipped with the $\sigma-$algebra $\mathcal{B}_l$ of Borel sets (the smallest $\sigma-$algebra containing all open (equivalently closed) subsets of X_l). Let μ_l be a positive, regular, Borel, probability measure on X_l, that is, a positive semi-definite, $\sigma-$additive function on $\mathcal{B}_l$ with $\mu_l(X_l) = 1$ and regularity means that the measure of every measurable set can be approximated arbitrarily well by open and compact sets (hence closed since X_l is compact Hausdorff) respectively. Since the measure is Borel, the continuous functions $C(X_l)$ are automatically measurable.

Definition I.2.17
A family of measures $(\mu_l)_{l\in\mathcal{L}}$ on the projections X_l of a projective family $(X_l, p_{ll'})_{l\prec l'\in\mathcal{L}}$ where the $p_{l'l}: X_{l'} \to X_l$ are continuous and surjective projections is said to be consistent provided that

$$(p_{l'l})_* \mu_{l'} := \mu_{l'} \circ p_{l'l}^{-1} = \mu_l \tag{I.2.3. 1}$$

for any $l \prec l'$. The measure $(p_{l'l})_*\mu_{l'}$ on X_l is called the push-forward of the measure $\mu_{l'}$.

The meaning of condition (I.2.3. 1) is the following: Let $\mathcal{B}_l \ni U_l \subset X_l$ be measurable. Since $p_{l'l}$ is continuous the pre-images of open sets in X_l are open in $X_{l'}$ and therefore measurable, hence $p_{l'l}$ is measurable. Since U_l is generated from countable unions and intersections of open sets it follows that $p_{l'l}^{-1}(U_l)$ is measurable. Then we require that

$$\mu_{l'}(p_{l'l}^{-1}(U_l)) = \mu_l(U_l) \tag{I.2.3. 2}$$

for any measurable U_l. We can rewrite condition (I.2.3. 2) in the form

$$\int_{X_{l'}} d\mu_{l'}(x_{l'}) \chi_{p_{l'l}^{-1}(U_l)}(x_{l'}) = \int_{X_l} d\mu_l(x_l) \chi_{U_l}(x_l) \tag{I.2.3. 3}$$

where χ_S denotes the characteristic function of a set S. Here it is strongly motivated to have surjective projections $p_{l'l}$ as otherwise $p_{l'l}^{-1}(X_l)$ is a proper subset of $X_{l'}$ so that $1 = \mu_l(X_l) = \mu_{l'}(p_{l'l}^{-1}(X_l)$ could give a contradiction with the μ_l being probability measures if $X_{l'} - p_{l'l}^{-1}(X_l)$ is not a set of measure zero with respect to $\mu_{l'}$.

Condition (I.2.3. 3) extends linearly to linear combinations of characteristic functions, so-called simple functions (see section III.5) and the (Lebesgue) integral of any measurable function is defined in terms of simple functions (see section III.5). Therefore we may equivalently write (I.2.3. 1) as

$$\int_{X_{l'}} d\mu_{l'}(x_{l'})[p_{l'l}^* f_l](x_{l'}) = \int_{X_l} d\mu_l(x_l) f_l(x_l) \tag{I.2.3. 4}$$

for any $l \prec l'$ and any $f_l \in C(X_l)$ since every measurable function can be approximated by simple functions and measurable simple functions can be approximated by continuous functions (which are automatically measurable). In the form (I.2.3. 4) the consistency condition means that integrating out the degrees of freedom in $X_{l'}$ on which $p_{l'l}^* f_l$ does not depend, we end up with with the same integral as if we had integrated over X_l only.

To summarize:
Let $f = [f_l]_\sim \in \mathrm{Cyl}(\overline{X})$ with $f_l \in C(X_l)$. Then (I.2.3. 4) ensures that the linear functional

$$\Lambda:\ \mathrm{Cyl}(\overline{X}) \to \mathbb{C};\ f = [f_l]_\sim \mapsto \Lambda(f) := \int_{X_l} d\mu_l(x_l) f_l(x_l) \tag{I.2.3. 5}$$

is well defined, i.e independent of the representative $f_l \sim p_{l'l}^* f_l$ of f. Moreover, it is a positive linear functional (integrals of positive functions are positive) because the μ_l are positive measures. Since $\mathrm{Cyl}(\overline{X}) \subset \overline{\mathrm{Cyl}(\overline{X})}$ is a subset of a unital C^*-algebra, Λ is automatically continuous (see the end of section III.5) and therefore extends uniquely and continuously to the completion $\overline{\mathrm{Cyl}(\overline{X})}$ by the bounded linear transformation theorem. Now in sections I.2.1.2, I.2.1.3 we showed that the Gel'fand isomorphism applied to $\overline{\mathrm{Cyl}(\overline{X})}$ leads to an (isometric) isomorphism of $\overline{\mathrm{Cyl}(\overline{X})}$ with $C(\overline{X})$ given by

$$\bigvee:\ \mathrm{Cyl}(\overline{X}) \to C(\overline{X});\ f = [f_l]_\sim \mapsto p_l^* f_l \tag{I.2.3. 6}$$

(and extended to $\overline{\mathrm{Cyl}(\overline{X})}$ using that $\mathrm{Cyl}(\overline{X})$ is dense). It follows that we may consider (I.2.3. 4) as a positive linear functional on $C(\overline{X})$. Since $\overline{X}$ is a compact Hausdorff space we are in position to apply the Riesz representation theorem.

Theorem I.2.10
Let $(X_l, p_{l'l})_{l\prec l'\in\mathcal{L}}$ be a compact Hausdorff projective family with continuous and surjective projections $p_{l'l}:\ X_{l'} \to X_l$, projective limit $\overline{X}$ and projections $p_l:\ \overline{X} \to X_l$.
i)
If μ is a regular Borel probability measure on $\overline{X}$ then $(\mu_l := \mu \circ p_l^{-1})_{l\in\mathcal{L}}$ defines a consistent family of regular Borel probability measures on X_l.
ii)
If $(\mu_l)_{l\in\mathcal{L}}$ defines a consistent family of regular Borel probability measures on X_l then there exists a unique, regular Borel probability measure μ on $\overline{X}$ such that $\mu \circ p_l^{-1} = \mu_l$.
iii)
The measure μ is faithful if and only if every μ_l is faithful.

Proof of Theorem I.2.10:
i)
Define the following positive lineal functional on $C(X_l)$:

$$\Lambda_l:\ C(X_l) \to \mathbb{C};\ f_l \mapsto \int_{\overline{X}} d\mu(x)(p_l^* f_l)(x) \tag{I.2.3. 7}$$

which satisfies $\Lambda_l(1) = 1$. Since X_l is a compact Hausdorff space, by the Riez representation theorem there exists a unique, positive, regular Borel probability measure μ_l on X_l that represents Λ_l, that is

$$\Lambda_l(f_l) = \int_{X_l} d\mu_l(x_l) f_l(x_l) \tag{I.2.3. 8}$$

Since $p_{l'l} \circ p_{l'} = p_l$, the consistency condition (I.2.3. 4) is obviously met.
ii)
As was shown above, the positive linear functional on $C(\overline{X})$

$$\Lambda: \ C(\overline{X}) \to \mathbb{C}; \ f = p_l^* f_l \equiv [f_l]_\sim \mapsto \int_{X_l} d\mu_l(x_l) f_l(x_l) \tag{I.2.3. 9}$$

is well-defined due to the consistency condition and satisfies $\Lambda(1) = 1$. Since $\overline{X}$ is a compact Hausdorff space the Riesz representation theorem guarantees the existence of a unique, positive, regular Borel probability measure μ on $\overline{X}$ representing Λ, that is

$$\Lambda(f) = \int_{\overline{X}} d\mu(x) f(x) \tag{I.2.3. 10}$$

iii)
Consider $f \in C(\overline{X})$ of the form $f = p_l^* f_l$ for some $l \in \mathcal{L}, \ f_l \in C(X_l)$. Functions of the form $p_l^* f_l$ lie dense in $C(\overline{X})$. Now $f = p_l^* f_l$ is non-negative iff f_l is non-negative because p_l is a surjection. It follows that we can restrict attention to all non-negative functions of the form $f = p_l^* f_l$ for arbitrary $f_l \in C(X_l), \ l \in \mathcal{L}$ as far as faithfulness is concerned. Let $\Lambda_\mu, \Lambda_{\mu_l}$ be the positive linear functionals determined by μ, μ_l respectively. Then:
μ faithful $\Leftrightarrow$ $\Lambda_\mu(p_l^* f_l) = \Lambda_{\mu_l}(f_l) = 0$ for any non-negative $f_l \in C(X_l)$ and any $l \in \mathcal{L}$ implies $f = 0$ $\Leftrightarrow$ For any $l \in \mathcal{L}$ and any non-negative $f_l \in C(X_l)$ the condition $\Lambda_{\mu_l}(f_l)$ implies $f_l = 0 \Leftrightarrow$ all μ_l are faithful.
□

We now define a natural measure on the spectrum of interest namely $\overline{\mathcal{A}}$, the so-called uniform measure. To do this we must specify the space of cylindrical functions. Given a subgroupoid $l \in \mathcal{L}$ with $l = l(\gamma)$ we think of an element $x_l \in X_l$ as a collection of group elements $\{x_l(e)\}_{e \in E(\gamma)} = \rho_l(x_l)$ and X_l can be identified with $G^{|E(\gamma)|}$ (see I.2.1. 16). Thus, an element $f_l \in C(X_l)$ is simply given by

$$f_l(x_l) = F_l(\{x_l(e)\}_{e \in E(\gamma)}) = (\rho_l^* F_l)(x_l) \tag{I.2.3. 11}$$

where F_l is a continuous complex valued function on $G^{|E(\gamma)|}$. For $l \prec l'$ with $l = l(\gamma), l' = l(\gamma')$ we define $\rho_{l'l}: \ G^{|E(\gamma')|} \to G^{|E(\gamma)|}$ by $\rho_l \circ p_{l'l} = \rho_{l'l} \circ \rho_{l'}$ (recall that ρ_l is a bijection.

Definition I.2.18
Let $\mathcal{L}$ be the set of all tame subgroupoids of the set of piecewise analytic paths $\mathcal{P}$ in σ and $X_l = Hom(l, G)$ identified with $G^{|E(\gamma)|}$ if $l = l(\gamma)$ via (I.2.1. 16). Then we define for any $f \in C(X_l)$

$$\mu_{0l}(f_l) = \int_{X_l} d\mu_{0l}(x_l) \rho_l^* F_l(x_l) := \int_{G^{|E(\gamma)|}} [\prod_{e \in E(\gamma)} d\mu_H(h_e)] \ F_l(\{h_e\}_{e \in E(\gamma)}) \tag{I.2.3. 12}$$

where μ_H is the Haar probability measure on G which thanks to the comapctness of G is invariant under left – and right translations and under inversions.

Lemma I.2.9
The linear functionals μ_l in (I.2.3. 12) are positive and consistently defined.

Proof of Lemma I.2.9:
That μ_l defines a positive linear functional follows from the explicit formula (I.2.3. 11) in terms of the positive Haar measure on G^n. That $(\mu_{0l})_{l\in\mathcal{L}}$ defines a consistent family follows from the observation that if $l \prec l'$ with $l = l(\gamma), l' = l(\gamma')$ then we can reach l from l' by a finite combination of the following three steps:
a) $e_0 \in E(\gamma')$ but $e_0 \cap \gamma \subset \{b(e_0), f(e_0)\}$ (deletion of an edge).
b) $e_0 \in E(\gamma')$ but $e_0^{-1} \in E(\gamma)$ (inversion of an edge).
c) $e_1, e_2 \in E(\gamma')$ but $e_0 = e_1 \circ e_2 \in E(\gamma)$ (composition of edges)
It therefore suffices to establish consistency with respect to all of these elementary steps.

In general we have

$$p_{l'l}^* f_l = p_{l'l}^* \rho_l^* F_l = \rho_{l'}^* \rho_{l'l}^* F_l \tag{I.2.3. 13}$$

whence

$$\mu_{0l'}(p_{l'l}^* f_l) = \mu_{0l'}(\rho_{l'}^*[\rho_{l'l}^* F_l]) = \int_{G^{|E(\gamma')|}} [\prod_{e\in E(\gamma')} d\mu_H(h_e)]\ [\rho_{l'l}^* F_l](\{h_e\}_{e\in E(\gamma')}) \tag{I.2.3. 14}$$

In what follows we will interchange freely orders of integration and break the integral over G^n in integrals over G^m, G^{n-m}. This is allowed by Fubini's theorem since the integrand, being bounded, is absolutely integrable in any order.
a)
We have $\rho_{l'l}(\{h_e\}_{e\in E(\gamma')}) = \{h_e\}_{e\in E(\gamma)}$ thus

$$\mu_{0l'}(p_{l'l}^* f_l) = \{\int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(h_e)] F_l(\{h_e\}_{e\in E(\gamma)})\}\{\int_G d\mu_H(h_{e_0})\ 1\} = \mu_{0l}(f_l) \tag{I.2.3. 15}$$

since μ_H is a probability measure.
b)
We have $\rho_{l'l}(\{h_e\}_{e\in E(\gamma')}) = \{\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_0}^{-1}\}$ thus

$$\begin{aligned}
\mu_{0l'}(p_{l'l}^* f_l) &= \int_{G^{|E(\gamma)|-1}} [\prod_{e\in E(\gamma)-\{e_0\}} d\mu_H(h_e)] \int_G d\mu_H(h_{e_0}) F_l(\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_0^{-1}}) \\
&= \int_{G^{|E(\gamma)-1|}} [\prod_{e\in E(\gamma)-\{e_0\}} d\mu_H(h_e)] \int_G d\mu_H(h_{e_0}^{-1}) F_l(\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_0}^{-1}) \\
&= \int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(h_e)] F_l(\{h_e\}_{e\in E(\gamma)}) = \mu_{0l}(f_l)
\end{aligned} \tag{I.2.3. 16}$$

since the Jacobian of the Haar measure with respect to the inversion map on G equals unity and where we have defined a new integration variable $h_{e_0^{-1}} := h_{e_0}^{-1}$.
c)
We have $\rho_{l'l}(\{h_e\}_{e\in E(\gamma')}) = \{\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_1} h_{e_2}\}$ thus

$$\begin{aligned}
&\mu_{0l'}(p_{l'l}^* f_l) \\
= &\int_{G^{|E(\gamma)|-1}} [\prod_{e\in E(\gamma)-\{e_0\}} d\mu_H(h_e)] \int_{G^2} d\mu_H(h_{e_1}) d\mu_H(h_{e_2}) F_l(\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_1} h_{e_2})
\end{aligned}$$

$$
\begin{aligned}
&= \int_{G^{|E(\gamma)|-1}}[\prod_{e\in E(\gamma)-\{e_0\}} d\mu_H(h_e)]\int_G d\mu_H(h_{e_1})\int_G d\mu_H(h_{e_1}^{-1}h_{e_1\circ e_2})\times \\
&\quad \times F_l(\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_1\circ e_2}) \\
&= \int_{G^{|E(\gamma)|-1}}[\prod_{e\in E(\gamma)-\{e_0\}} d\mu_H(h_e)]\int_G d\mu_H(h_{e_1\circ e_2})F_l(\{h_e\}_{e\in E(\gamma)-\{e_0\}}, h_{e_1\circ e_2})\times \\
&\quad \times[\int_G d\mu_H(h_{e_1})1] \\
&= \int_{G^{|E(\gamma)|}}[\prod_{e\in E(\gamma)} d\mu_H(h_e)]F_l(\{h_e\}_{e\in E(\gamma)}) = \mu_{0l}(f_l) \qquad \text{(I.2.3. 17)}
\end{aligned}
$$

since the Jacobian of the Haar measure with respect to the left or right translation map on G equals unity and where we have defined a new integration variable by $h_{e_1\circ e_2} := h_{e_1}h_{e_2}$.
□
It follows from theorem I.2.10 that the family (μ_{0l}) defines a regular Borel probability measure on $\overline{X}$.

We now can equip the quantum configuration space $\overline{\mathcal{A}}$ with a Hilbert space structure.

Definition I.2.19
The Hilbert space $\mathcal{H}^0$ is defined as the space of square integrable functions over $\overline{\mathcal{A}}$ with respect to the uniform measure μ_0, that is

$$\mathcal{H}_0 := L_2(\overline{\mathcal{A}}, d\mu_0) \qquad \text{(I.2.3. 18)}$$

Notice that since we have identified cylindrical functions over $\overline{\mathcal{A}/\mathcal{G}}$ with gauge invariant, cylindrical functions over $\overline{\mathcal{A}}$ the measure μ_0 can also be defined as a measure on $\overline{\mathcal{A}/\mathcal{G}}$: Simply restrict the μ_{0l} to the invariant elements which still defines a positive linear functional on $C([X_l]_l)$ and then use the Riez representation theorem. It is easy to check that the obtained measure coincides with the restriction of μ_0 to $\overline{\mathcal{A}/\mathcal{G}}$ with $\sigma-$algebra given by the sets $U\cap\overline{\mathcal{A}/\mathcal{G}}$ where U is measurable in $\overline{\mathcal{A}}$. We will denote the restricted and unrestricted measure by the same symbol μ_0.

At this point the physical significance of the Hilbert space is unclear because we did not show that it supports a representation of the canonical commutation relations and adjointness relations. We will demonstrate this to be the case in section I.3.

I.2.4 Functional Calculus on a Projective Limit

This section rests on the simple but powerful obeservation that in the case of interest the projections $p_{l'l}$ are not only continuous and surjective but also analytic. This can be seen by using the bijection (I.2.1. 16) between X_l and G^n for some n and using the standard differentiable structure on G^n.

Functions
We have seen that we can identify $C(\overline{X})$ with the (completion of the) space of) cylindrical functions $f = [f_l]/\sim = p_l^* f_l,\ f_l \in C(X_l)$. This suggests to proceed analogously with the other differentiability categories. Let $n \in \{0, 1, 2, ..\} \cup \{\infty\} \cup \{\omega\}$ then we define

$$\mathrm{Cyl}^n(\overline{X}) := (\bigcup_{l\in\mathcal{L}} C^n(X_l))/\sim \qquad \text{(I.2.4. 1)}$$

That is, a typical element $f = [f_l]_\sim \in \mathrm{Cyl}^n(\overline{X})$ can be thought of as an equivalence class of elements of the form $f_l \in C^n(X_l)$ where $f_l \sim f_{l'}$ iff there exists $l, l' \prec l''$ such that $p_{l''l}^* f_l = p_{l''l'}^* f_{l'}$. As in the previous section, the existence of one such l'' implies that this equation holds for all $l, l' \prec l''$. Notice that $f_l \in C^n(X_l)$ implies $p^*_{l'l} f_l \in C^n(X_l)$ due to the analyticity of the projections, this is where their analyticity becomes important.

Differential Forms
In fact, since the Grassman algebra of differential forms on X_l is generated by finite linear combinations of monomials of the form $f_l^{(0)}\ df_l^{1)} \wedge .. \wedge df_l^{(p)}$ with $0 \le p \le \dim(X_l)$, $f_l^{(0)} \in C^n(X_l)$, $f_l^{(k)} \in C^{(n+1)}(X_l)$, $k = 1, .., p$ we can define the space of cylindrical $p-$forms and the cylindrical Grassman algebra by

$$\bigwedge^n(\overline{X}) = (\bigcup_{l \in \mathcal{L}} \bigwedge^n(X_l))/\sim \qquad \text{(I.2.4. 2)}$$

because the pull-back commutes with the exterior derivative, that is, $p^*_{l''l} f_l = p_{l''l'} f_{l'}$ implies $p^*_{l''l} df_l = d(p_{l''l'} f_{l'})$, in other words, the exterior derivative is a well-defined operation on the Grassmann algebra. Notice that if $\omega = [\omega_l]_\sim \in \bigwedge(\overline{X})$ and ω_l has degree p then also $p^*_{l'l}\omega_l$ has degree p, hence the degree of forms on $\overline{X}$ is well-defined.

Volume Forms
The case of volume forms is slightly different because a volume form on an orientable X_l is a nowehere vanishing differential form of degree $\dim(X_l)$ so that the degree varies with the label l. However, volume forms on $\overline{X}$ (even in the non-orientable case) are nothing else than cylindrically defined measures satisfying the consistency condition $\mu_{l'} \circ p_{l'l} = \mu_l$ for all $l \prec l'$. If they are probability measures we can extend them to $\sigma-$additive measures on $\overline{X}$ using the Riesz-Markow theorem as in the previous section.

Vector Fields
Differentiable vector fields $V^n(X_l)$ on X_l are conveniently introduced algebraically on X_l as derivatives, that is, they are linear functionals $Y_l;\ C^{n+1}(X_l) \to C^n(X_l)$ annihilating constants and satisfying the Leibniz rule. We want to proceed similarly with respect to $\overline{X}$ and the first impulse would be to define

$$V^n(\overline{X}) = (\bigcup_{l \in \mathcal{L}} V^n(X_l))/\sim$$

where the equivalence relation is given through the push-forward map. The push-forward is defined by

$$(p_{l'l})_* :\ V^n(X_{l'}) \to V^n(X_{l'}); p^*_{l'l}([(p_{l'l})_* Y_{l'}](f_l)) := Y_{l'}(p^*_{l'l} f_l) \qquad \text{(I.2.4. 3)}$$

and we could try to define $Y_l \sim Y_{l'}$ iff for any $l'' \prec l, l'$ we have $(p_{l'l''})_* Y_{l'} = (p_{ll''})_* Y_l$. The problem with this definition is that the push-forward moves us "down" in the directed label set $\mathcal{L}$ instead of "up" as is the case with the pull-back so that it is not guaranteed that, given l, l' there exists any l'' at all that satisfies $l, l' \prec l''$ whence the consistency condition might be empty. This forces us to adopt a different strategy, namely to define $V^n(\overline{X})$ as projective nets $(Y_l)_{l_0 \prec l \in \mathcal{L}}$ with the consistency condition

$$(p_{l'l})_* Y_{l'} = Y_l \Leftrightarrow p^*_{l'l}[Y_l(f_l)] = Y_{l'}(p^*_{l'l} f_l)\ \forall\ f_l \in C^n(X_l)\ l_0 \prec l \prec l' \qquad \text{(I.2.4. 4)}$$

The necessity to restrict attention to $l_0 \prec l$ is that it may not be possible or necessary to define Y_l for all $l \in \mathcal{L}$ or to have (I.2.4. 4) satisfied. This question never came up of course for the pull-back. Notice that (I.2.4. 4) means that if $f_{l'} = p^*_{l'l} f_l$ then $Y_{l'}(f_{l'}) = p^*_{l'l} Y_l(f_l)$ for $l_0 \prec l \prec l'$, that is consistently defined vector fields map cylindrical fucntions to cylindrical functions.

It is clear that for $f=[f_l]_\sim=p_l^* f_l$ with $l_0\prec l$ the formula

$$Y(p_l^* f_l):=p_l^* Y_l(f_l)=:p_l^*[(p_l)_* Y](f_l) \tag{I.2.4. 5}$$

is well-defined for suppose that $f_l\sim f_l'$ with $l_0\prec l'$ then we find $l_0\prec l,l'\prec l''$ such that $p_{l''l}^* f_l=p_{l''l'}^* f_{l'}$ whence, using $p_{l''l}\circ p_{l''}=p_l, p_{l''l'}\circ p_{l''}=p_{l'}$

$$p_{l'}^* Y_{l'}(f_{l'})=p_{l''}^* p_{l''l'}^* Y_{l'}(f_{l'})=p_{l''}^* Y_{l''}(p_{l''l'}^* f_{l'})=p_{l''}^* Y_{l''}(p_{l''l}^* f_l)=p_l^* Y_l(f_l) \tag{I.2.4. 6}$$

Lie Brackets

Suppose that $Y=(Y_l)_{l_0\prec l\in\mathcal{L}}, Y'=(Y_l')_{l_0'\prec l\in\mathcal{L}}\in V^n(\overline{X})$ are consistently defined vector fields. We certainly find $l_0,l_0'\prec l_0''$ and claim that $[Y,Y']:=([Y_l,Y_l])_{l_0''l\in\mathcal{L}}\in V^{n-1}(\overline{X})$ is again consistently defined. To see this, consider $l_0^{\prec}l\prec l'$ then for any $f_l\in C^n(X_l)$ we have due to $l_0\prec l$ and $l_0'\prec l$

$$p_{l'l}^*([Y_l,Y_l'](f_l))=Y_{l'}[p_{l'l}^*(Y_l'(f_l))]-Y_{l'}'[p_{l'l}^*(Y_l(f_l))]=[Y_{l'},Y_{l'}'](p_{l'l}^* f_l) \tag{I.2.4. 7}$$

Vector Field Divergences

Recall that the Lie derivative of an element $\omega_l\in\bigwedge^n(X_l)$ with respect to a vector field $Y_l\in V^n(X_l)$ is defined by $L_{Y_l}\omega_l=[i_{Y_l}d+di_{Y_l}]\omega_l$ where

$$i_{Y_l}f_l^{(0)}\,df_l^{(1)}\wedge..\wedge df_l^{(p)}=f_l^{(0)}\sum_{k=1}^p(-1)^{k+1}Y_l(f_l^{(k)})\,df_l^{(1)}\wedge..df_l^{(k-1)}\wedge df_l^{(k+1)}\wedge..\wedge df_l^{(p)}$$

denotes contraction of forms with vector fields, annihilating zero forms. Let now μ_l be a volume form on X_l. Since X_l is finite dimesional, all smooth volume forms are absolutely continuous with respect to each other and there exists a well-defined function, called the divergence of Y_l with respect to μ_l, uniquely defined by

$$L_{Y_l}\mu_l=:[\mathrm{div}_{\mu_l}Y_l]\mu_l \tag{I.2.4. 8}$$

We say that a vector field $Y=(Y_l)_{l_0\prec l\in\mathcal{L}}$ is compatible with a volume form $\mu=(\mu_l)_{l\in\mathcal{L}}$ provided that the family of divergences defines a cylindrical function, that is

$$p_{l'l}^*[\mathrm{div}_{\mu_l}Y_l]=\mathrm{div}_{\mu_{l'}}Y_{l'}\ \forall l_0\prec l\prec l' \tag{I.2.4. 9}$$

Hence there exists a well defined cylindrical function $\mathrm{div}_\mu Y:=[\mathrm{div}_{\mu_l}Y_l]_\sim$, called the divergence of Y with respect to μ.

Lemma I.2.10 *Let μ be a smooth volume form, Y,Y' $\mu-$compatible vector fields and $f,f'\in Cyl^1(\overline{X})$ cylindrical functions on $\overline{X}$.*
i)
If $\partial X_l=\emptyset$ has no boundary then

$$\int_{\overline{X}}\mu\ f\ Y(f')=-\int_{\overline{X}}\mu\ (Y(f)+f\ [div_\mu Y])f' \tag{I.2.4. 10}$$

ii)
The Lie bracket $[Y,Y']$ is again $\mu-$compatible and

$$div_\mu[Y,Y']=Y(div_\mu Y')-Y'(div_\mu Y) \tag{I.2.4. 11}$$

Proof of Lemma I.2.10:
i)
We find $l_0, l_0' \prec l$ such that $f = p_l^* f_l, f' = p_l^* f_l'$. Then

$$\begin{aligned}
\mu(fY(f')) &= \mu([p_l^* f_l][p_l^* Y_l(f_l')]) = \mu_l(f_l L_{Y_l}[f_l']) = \int_{X_l} \{L_{Y_l}[\mu_l f_l f_l'] - (L_{Y_l}[\mu_l f_l]) f_l'\} \\
&= \int_{X_l} \{d\ i_{Y_l}[\mu_l f_l f_l'] - \mu_l(Y_l(f_l) + f_l\ [\mathrm{div}_{\mu_l} Y_l]) f_l'\} \\
&= -\mu((Y(f) + f\ [\mathrm{div}_\mu Y]) f') \qquad \text{(I.2.4. 12)}
\end{aligned}$$

where in the third line we have applied Stokes' theorem and that the Lie derivative satisfies the Leibniz rule.
ii)
We find $l_0, l_0' \prec l_0''$ so that $([Y_l, Y_l'])_{l_0'' \prec l \in \mathcal{L}}$ is consistently defined as shown above. From the fact that the Lie derivative is an isomorphism between the Lie algebra of vector fields and the derivatives respectively on $C^n(X_l)$, $L_{[Y_l, Y_l']} = [L_{Y_l}, L_{Y_l'}]$, and the fact that Lie derivation and exterior derivation commute, $[d, L_{Y_l}] = 0$, we have

$$\begin{aligned}
(\mathrm{div}_{\mu_l}[Y_l, Y_l'])\mu_l &= [L_{Y_l}, L_{Y_l'}](\mu_l) = L_{Y_l}([\mathrm{div}_{\mu_l} Y_l']\mu_l) - L_{Y_l'}([\mathrm{div}_{\mu_l} Y_l]\mu_l) \\
&= [Y_l(\mathrm{div}_{\mu_l} Y_l') - Y_l'(\mathrm{div}_{\mu_l} Y_l)]\mu_l \qquad \text{(I.2.4. 13)}
\end{aligned}$$

It follows from the consistency of the Y_l and the compatibility with the μ_l that for $l \prec l'$

$$p_{ll'}^* Y_l(\mathrm{div}_{\mu_l} Y_l') = Y_{l'}(p_{ll'}^*(\mathrm{div}_{\mu_l} Y_l')) = Y_{l'}(\mathrm{div}_{\mu_{l'}} Y_{l'}') \qquad \text{(I.2.4. 14)}$$

□

Momentum Operators
Let Y be a vector field compatible with $\sigma-$additive measure (volume form) μ such that it is together with its divergence $\mathrm{div}_\mu Y$ is real valued. We consider the Hilbert space $\mathcal{H}_\mu := L_2(\overline{X}, \mu)$ and define the momentum operator

$$P(Y) := i(Y + \frac{1}{2}(\mathrm{div}_\mu Y)1_{\mathcal{H}_\mu}) \qquad \text{(I.2.4. 15)}$$

with dense domain $D(P(Y)) = \mathrm{Cyl}^1(\overline{X})$. From (I.2.4. 10) we conclude that for $f, f' \in D(P(Y))$

$$< f, P(Y)f' >_\mu = \mu(\overline{f} P(Y) f) = \mu(\overline{P(Y) f} f) = < P(Y)f, f' >_\mu \qquad \text{(I.2.4. 16)}$$

from which we see that

$$D(P(Y)) \subset D(P(Y)^\dagger) := \{f \in \mathcal{H}_\mu;\ \sup_{||f'||>0} | < f, P(Y)f' > |/||f'|| < \infty \text{ and } D(P(Y))^\dagger_{|D(P(Y))} = P(Y)$$

whence $P(Y)$ is a symmetric unbounded operator.
Finally we notice that if Y, Y' are both $\mu-$compatible then

$$[P(Y), P(Y')] = iP([Y, Y']) \qquad \text{(I.2.4. 17)}$$

by a straightforward computation using lemma I.2.10.

Remark:
That $\mathrm{div}_{\mu_l} Y_l$ is a cylindrical function is a sufficient criterion for $P(Y)$ to be well defined, but it

is too strong a requirement because it means that for given l on any other $l \prec l'$ the function $div_{\mu_{l'}} Y_{l'} \equiv p^*_{l'l}(div_{\mu_l} Y_l)$ does not depend on the additional degrees of freedom contained in $X_{l'}$. That is, if not some special graphs are too be distinguished then $\mathrm{div}_\mu Y = const.$ is the only possibility. So compatibility between μ and Y is only sufficient but has not been shown to be necessary in order to define interesting momentum operators. It would be important to replace the compatibility criterion by a weaker one.

General Operators
More generally we have the following abstract situation:
a)
We have a partially ordered and directed index set $\mathcal{L}$, a family of Hilbert spaces $\mathcal{H}_l := \mathcal{H}_{\mu_l} := L_2(X_l, d\mu_l)$ and isometric monomorphisms (linear injections)

$$\hat{U}_{ll'} : \mathcal{H}_l \to \mathcal{H}_{l'} \tag{I.2.4. 18}$$

for every $l \prec l'$ which in our special case is given by $\hat{U}_l f_l := p^*_{l'l} f_l$. The isometric monomorphisms satisfy the compatibility condition

$$\hat{U}_{l'l''} \hat{U}_{ll'} = \hat{U}_{ll''} \tag{I.2.4. 19}$$

for any $l \prec l' \prec l''$ due to $p_{l'l} \circ p_{l''l'} = p_{l''l}$. A system $(\mathcal{H}_l, \hat{U}_{ll'})_{l \prec l' \in \mathcal{L}}$ of this sort is called a directed system of Hilbert spaces. A Hilbert space $\mathcal{H}$ is called the inductive limit of a directed system of Hilbert spaces provided that there exist isometric monomorphisms

$$\hat{U}_l : \mathcal{H}_l \to \mathcal{H} \tag{I.2.4. 20}$$

for any $l \in \mathcal{L}$ such that the compatibility condition

$$\hat{U}_{l'} \hat{U}_{ll'} = \hat{U}_l \tag{I.2.4. 21}$$

holds. In our case, obviously $\hat{U}_l f_l := p^*_l f_l$ provides these monomorphisms so that we have displayed $\mathcal{H}_\mu$ as the inductive limit of the $\mathcal{H}_{\mu_l}$.

Likewise we have a family of operators $\hat{O}_l = P(Y_l)$ with dense domain $D(\hat{O}_l) = C^1(X_l)$ in $\mathcal{H}_l$ which are defined for a cofinal subset $\mathcal{L}(\hat{O}) = \{l \in \mathcal{L};\ l_0 \prec l\}$ (that is, for any $l \in \mathcal{L}$ there exists $l \prec l' \in \mathcal{L}(\hat{O})$) of $\mathcal{L}$. These families of domains and operators satisfy the following compatibility conditions:

$$\hat{U}_{ll'} D(\hat{O}_l) \subset D(\hat{O}_{l'}) \tag{I.2.4. 22}$$

for any $l \prec l' \in \mathcal{L}(\hat{O})$ since $p^*_{l'l} C^1(X_l) \subset C^1(X_{l'})$ (the pull-back of functions is C^1 with respect to the X_l arguments but C^ω with respect to the remaining arguments in $X_{l'}$). Furthermore

$$\hat{U}_{ll'} \hat{O}_l = \hat{O}_{l'} \hat{U}_{ll'} \tag{I.2.4. 23}$$

for any $l \prec l' \in \mathcal{L}(\hat{O})$ since $p^*_{l'l}(Y_l(f_l) + [\mathrm{div}_{\mu_l} Y_l] f_l/2) = (Y_{l'}(p^*_{l'l} f_l) + [\mathrm{div}_{\mu_{l'}} Y_{l'}] p^*_{l'l} f_l/2)$ due to consistency and compatibility. A structure of this kind is called a directed system of operators. An operator $\hat{O}$ with dense domain $D(\hat{O})$ is called the inductive limit of a directed system of operators provided the above defined isometric isomorphisms interact with domains and operators in the expected way, that is,

$$\hat{U}_l D(\hat{O}_l) \subset D(\hat{O}) \tag{I.2.4. 24}$$

and

$$\hat{U}_l \hat{O}_l = \hat{O} \hat{U}_l \tag{I.2.4. 25}$$

In our case this is by definition statisfied since $p_l^* C^1(X_l) \subset \mathrm{Cyl}^1(\overline{X})$ and $p_l^*(Y_l(f_l) + [\mathrm{div}_{\mu_l} Y_l] f_l/2) \equiv (Y(p_l^* f_l) + [\mathrm{div}_\mu Y] p_l^* f_l/2)$.

It turns out that directed systems of Hilbert spaces and operators always have an inductive limit which is unique up to unitary equivalence.

Lemma I.2.11
i)
Given directed systems of Hilbert spaces $(\mathcal{H}_l, \hat{U}_{ll'})_{l \prec l' \in \mathcal{L}}$ and operators $(\hat{O}_l, D(\hat{O}_l), \hat{U}_{ll'})_{l \prec l' \in \mathcal{L}(\hat{O})}$ with a cofinal index set $\mathcal{L}(\hat{O})$, there is a, up to unitary equivalence, unique inductive limit Hilbert space $(\mathcal{H}, \hat{U}_l)_{l \in \mathcal{L}}$ as well as a unique inductive limit operator $(\hat{O}, D(\hat{O}), \hat{U}_l)_{l \in \mathcal{L}(\hat{O})}$ densely defined on the inductive limit Hilbert space.
ii)
If the $\hat{O}_l$ are essentially self-adjoint with core $D(\hat{O}_l)$ then $\hat{O}$ is essentially self-adjoint with core $D(\hat{O})$.
iii)
If the $\hat{O}_l$ are essentially self-adjoint then $(\hat{O}_l', D(\hat{O}_l'), \hat{U}_{ll'})_{l \prec l' \in \mathcal{L}(\hat{O})}$ is a directed system of operators where O_l' denotes the self-adjoint extension of $\hat{O}_l$.

Proof of Lemma I.2.11:
i)
In the case of bounded operators, that is $D(\hat{O}_l) = \mathcal{H}_l$, part i) is standard in operator theory, see e.g. vol. 2 of the first reference of [142] for more details and an extension of the theorem to directed systems of C^*-algebras and von Neumann algebras which have a unique inductive limit up to algebra isomorphisms.

We consider the vector space V of equivalence classes of nets $f = (f_l)_{l_0 \prec l \in \mathcal{L}(f)}$ for some cofinal $\mathcal{L}(f) \subset \mathcal{L}$ with $f_l \in \mathcal{H}_l$ satisfying $\hat{U}_{ll'} f_l = f_{l'}$ for any $l_0 \prec l \prec l'$ and where $f \sim f'$ are equivalent if $f_l = f_l'$ for all $l \in \mathcal{L}(f) \cap \mathcal{L}(f')$. Let us write $[f]_\sim$ for the equivalence class of f. We define

$$\hat{U}_l : \ \mathcal{H}_l \to V; \ f_l \mapsto [(\hat{U}_{ll'} f_l)_{l \prec l' \in \mathcal{L}}]_\sim \tag{I.2.4. 26}$$

Due to isometry of the $\hat{U}_{ll'}$ the norm on V given by $||[f]_\sim|| := ||f_l||_l$ is independent of the choice of $l \in \mathcal{L}(f)$, in particular, $\hat{U}_l$ becomes an isometry. We have for $l \prec l'$

$$\hat{U}_{l'} \hat{U}_{ll'} f_l = [(\hat{U}_{l'l''} \hat{U}_{ll'} f_l)_{l' \prec l''}]_\sim = [(\hat{U}_{ll''} f_l)_{l' \prec l''}]_\sim = [(\hat{U}_{ll''} f_l)_{l \prec l''}]_\sim = \hat{U}_l f_l$$

Finally we consider the subspace of V given by the span of elements of the form $\hat{U}_l f_l$ with $f_l \in \mathcal{H}_l$ and complete it to arrive at the Hilbert space

$$\mathcal{H} := \overline{\bigcup_l \hat{U}_l \mathcal{H}_l} \tag{I.2.4. 27}$$

to which $\hat{U}_l$ can be extended uniquely as an isometric monomorphism by continuity. To see the uniqueness one observes that given another inductive limit $(\mathcal{H}', \hat{V}_l)$ we may define $W_l := \hat{V}_l \hat{U}_l^{-1} : \hat{U}_l \mathcal{H}_l \to \hat{V}_l \mathcal{H}_l$ which one checks to be an isometry. Also for $l \prec l'$ we have $W_{l'} \hat{U}_l = \hat{W}_{l'} \hat{U}_{l'} \hat{U}_{ll'} = \hat{V}_{l'} \hat{U}_{ll'} = V_l = \hat{W}_l \hat{U}_l$, in other words, $W_{l'}$ is an extension of W_l for $l \prec l'$. This means that we have a densely defined isometry $\hat{W} : \bigcup \hat{U}_l \mathcal{H}_l \to \bigcup \hat{V}_l \mathcal{H}_l$ defined by $\hat{W}_{|\hat{U}_l \mathcal{H}_l} = W_l$ which extends by continuity uniquely to an isometry between the two Hilbert spaces.

Next, define an operator on the dense subspace of $\mathcal{H}$ given by $D(\hat{O}) := \bigcup_{l \in \mathcal{L}(\hat{O})} \hat{U}_l D(\hat{O}_l)$

$$\hat{O}[(f_l)_{l \in \mathcal{L}(\hat{O})}]_\sim := [(\hat{O}_l f_l)_{l \in \mathcal{L}(\hat{O})}]_\sim \tag{I.2.4. 28}$$

Since $\mathcal{L}(\hat{O}) \cap \{l' \in \mathcal{L};\ l \prec l'\} = \{l' \in \mathcal{L}(\hat{O});\ l \prec l'\}$ is cofinal we have

$$\begin{aligned}\hat{O}\hat{U}_l f_l &= \hat{O}[(\hat{U}_{ll'} f_l)_{l\prec l'\in\mathcal{L}}]_\sim = \hat{O}[(\hat{U}_{ll'} f_l)_{l\prec l'\in\mathcal{L}(\hat{O})}]_\sim = [(\hat{O}_{l'}\hat{U}_{ll'} f_l)_{l\prec l'\in\mathcal{L}(\hat{O})}]_\sim \\ &= [(\hat{U}_{ll'}\hat{O}_l f_l)_{l\prec l'\in\mathcal{L}(\hat{O})}]_\sim = [(\hat{U}_{ll'}\hat{O}_l f_l)_{l\prec l'\in\mathcal{L}}]_\sim \\ &= \hat{U}_l\hat{O}_l f_l \end{aligned} \tag{I.2.4. 29}$$

ii)
By the basic criterion of essential self-adjointness we know that $(\hat{O}_l \pm i\cdot 1_{\mathcal{H}_l})D(\hat{O}_l)$ is dense in $\mathcal{H}_l$. It follows that

$$\begin{aligned}(\hat{O} \pm i\cdot 1_{\mathcal{H}})D(\hat{O}) &= \bigcup_{l\in\mathcal{L}(\hat{O})} (\hat{O} \pm i\cdot 1_{\mathcal{H}})\hat{U}_l D(\hat{O}_l) \\ &= \bigcup_{l\in\mathcal{L}(\hat{O})} \hat{U}_l(\hat{O}_l \pm i\cdot 1_{\mathcal{H}_l})D(\hat{O}_l)\end{aligned} \tag{I.2.4. 30}$$

hence $(\hat{O} \pm i\cdot 1_{\mathcal{H}})D(\hat{O})$ is dense in $\mathcal{H}$ so that $\hat{O}$ is essentially self-adjoint by the basic criterion of essential self-adjointness.
iii)
Recall that the self-adjoint extension $\hat{O}'_l$ of an essentially self-adjoint operator $\hat{O}_l$ with core $D(\hat{O}_l)$ is unique and given by its closure, that is, the set $D(\hat{O}'_l)$ given by those $f_l \in \mathcal{H}_l$ such that $(f_l, \hat{O}_l f_l) \in \overline{\Gamma}_{\hat{O}_l}$, the closure in $\mathcal{H}_l \times \mathcal{H}_l$ of the graph $\Gamma_{\hat{O}_l} = \{(f_l, \hat{O}_l f_l);\ f_l \in D(\hat{O}_l)\}$ of $\hat{O}_l$ with respect to the norm $||(f_l, f'_l)||^2 = ||f_l||^2 + ||f'_l||^2$.

To see that $\hat{U}_{ll'} D(\hat{O}'_l) \subset D(\hat{O}'_{l'})$ we notice that $\hat{U}_{ll'} D(\hat{O}_l) \subset D(\hat{O}_{l'})$. Hence, the closure $D(\hat{O}'_{l'})$ of $D(\hat{O}_{l'})$ will contain the closure of $\hat{U}_{ll'} D(\hat{O}_l)$ which coincides with $\hat{U}_{ll'} D(\hat{O}'_l)$ because $\hat{U}_{ll'}$ is bounded.

To see that $\hat{U}_{ll'}\hat{O}'_l = \hat{O}'_{l'}\hat{U}_{ll'}$ holds on $D(\hat{O}'_l)$ we notice that $\hat{U}_{ll'}\hat{O}_l = \hat{O}_{l'}\hat{U}_{ll'}$ holds on $D(\hat{O}_l)$. Since $\hat{O}'_l, \hat{O}'_{l'}$ are just the extensions of $\hat{O}_l, \hat{O}_{l'}$ from $D(\hat{O}_l), D(\hat{O}_{l'})$ to $D(\hat{O}'_l), D(\hat{O}'_{l'})$ and since $\hat{U}_{ll'} D(\hat{O}'_l) \subset D(\hat{O}'_{l'})$ the claim follows.
□

Passage to the Quotient Space
Finally we consider the case of interest, namely the quotient space $\overline{\mathcal{A}/\mathcal{G}}$ projective limit. The significance of the result $\overline{\mathcal{A}}/\overline{\mathcal{G}} = \overline{\mathcal{A}/\mathcal{G}}$ is that we can identify cylindrical functions on $\overline{\mathcal{A}/\mathcal{G}}$ simply with $\overline{\mathcal{G}}-$invariant functions on $\overline{\mathcal{A}}$. More precisely, if $\lambda :\ \overline{\mathcal{G}} \times \overline{\mathcal{A}} \to \overline{\mathcal{A}};\ A \mapsto \lambda_g(A)$ is the $\overline{\mathcal{G}}-$action and $f \in \mathrm{Cyl}^n(\overline{\mathcal{A}})$ is $\overline{\mathcal{G}}-$invariant then we may define $\tilde{f} \in \mathrm{Cyl}^n(\overline{\mathcal{A}/\mathcal{G}})$ by $\tilde{f}([A]) := f(A) = f(\lambda_g(A))$ for all $g \in \overline{\mathcal{G}}$ where $[.] :\ \overline{\mathcal{A}} \to \overline{\mathcal{A}}/\overline{\mathcal{G}} \equiv \overline{\mathcal{A}/\mathcal{G}}$ denotes the quotient map. Thus we define zero forms on $\overline{\mathcal{A}/\mathcal{G}}$ as zero forms on $\overline{\mathcal{A}}$ which satisfy $f = \lambda_g^* f$ for any $g \in \overline{\mathcal{G}}$. Notice that this is possible for any differentiability category because the $\overline{\mathcal{G}}-$action is evidently not only continuous but even analytic !

Since pull-backs commute with exterior derivation we can likewise define the Grassman algebra $\bigwedge(\overline{\mathcal{A}/\mathcal{G}})$ as the subalgebra of $\bigwedge(\overline{\mathcal{A}})$ given by the $\overline{\mathcal{G}}-$invariant differential forms, that is, those that satisfy $\lambda_g^*\omega = \omega$ for all $g \in \overline{\mathcal{G}}$ (if f is $\overline{\mathcal{G}}-$invariant, so is df because $\lambda_g^* df = d\lambda_g^* f = df$).

Next, volume forms on $\overline{\mathcal{A}/\mathcal{G}}$ are just $\overline{\mathcal{G}}-$invariant volume forms on $\overline{\mathcal{A}}$, that is $(\lambda_g)_*\mu = \mu \circ \lambda_g^{-1} = \mu \circ \lambda_{g^{-1}} = \mu$ for all $g \in \overline{\mathcal{G}}$. Given any volume form μ on $\overline{\mathcal{A}}$ we may derive a measure $\overline{\mu}$ on $\overline{\mathcal{A}/\mathcal{G}}$ by $\overline{\mu}(f) := \mu(f)$ for all $\overline{\mathcal{G}}-$invariant functions f on $\overline{\mathcal{A}}$. If we denote the Haar probability measure on $\overline{\mathcal{G}} \equiv \prod_{x\in\sigma} G$ by μ_H then from $\mu(f) = \mu(\lambda_g^* f) = [(\lambda_g)_*\mu](f)$ for all $\overline{\mathcal{G}}-$invariant measurable functions we find

$$\overline{\mu}([A]) = \int_{\overline{\mathcal{G}}} \mu_H(g)\ [(\lambda_g)_*\mu](A) \tag{I.2.4. 31}$$

Finally, we define vector fields on $\overline{\mathcal{A}/\mathcal{G}}$ as $\overline{\mathcal{G}}-$invariant vector fields $\overline{\mathcal{A}}$, that is, those satisfying $(\lambda_g)_* Y = Y$ for all $g \in \overline{\mathcal{G}}$, more precisely, if $Y = (Y_l)_{l_0 \prec l}$ then

$$(\lambda_g^l)^*([(\lambda_g^l)_* Y_l](f_l)) := Y_l[(\lambda_g^l)^* f_l] = (\lambda_g^l)^*(Y_l(f_l)) \tag{I.2.4. 32}$$

for any $f_l \in C^n(\overline{\mathcal{A}})$ and $l_0 \prec l$.

I.2.5 Density and Support Properties of $\mathcal{A}, \mathcal{A}/\mathcal{G}$ with Respect to $\overline{\mathcal{A}}, \overline{\mathcal{A}/\mathcal{G}}$

In this section we will see that $\mathcal{A}$ lies topologically dense, but measure theoretically thin in $\overline{\mathcal{A}}$ (similar results apply to $\mathcal{A}/\mathcal{G}$ with respect to $\overline{\mathcal{A}/\mathcal{G}} = \overline{\mathcal{A}}/\overline{\mathcal{G}}$) with respect to the uniform measure μ_0. More precisely, there is a dense embedding (injective inclusion) $\mathcal{A} \to \overline{\mathcal{A}}$ but $\mathcal{A}$ is embedded into a measurable subset of $\overline{\mathcal{A}}$ of measure zero. The latter result demonstrates that the measure is concentrated on non-smooth (distributional) connections so that $\overline{\mathcal{A}}$ is indeed much larger than $\mathcal{A}$.

We have seen in section I.2.1.2 that every element $A \in \mathcal{A}$ defines an element of $\mathrm{Hom}(\mathcal{P}, G)$ and that this space can be identified with the projective limit $\overline{X} \equiv \overline{\mathcal{A}}$. Now via the C^*-algebraic framework we know that $\overline{\mathrm{Cyl}(\overline{X})}$ can be identified with $C(\overline{X})$ and the latter space of functions separates the points of $\overline{X}$ by the Stone-Weierstrass theorem since it is Hausdorff and compact. The question is whether the smaller set of functions $\mathrm{Cyl}(\overline{X})$ separates the smaller set of points $\mathcal{A}$. This is almost obvious and we will do it for $G = SU(N)$, other compact groups can be treated similarly: Let $A \neq A'$ be given then there exists a point $x \in \sigma$ such that $A(x) \neq A'(x)$. Take $D = \dim(\sigma)$ edges $e_{x,\alpha} \in \mathcal{P}$ with $b(e_{x,\alpha}) = x$ and linearly independent tangents $\dot{e}_{x,\alpha}(0)$ at x. Consider the cylindrical function

$$F_x^\epsilon : \mathcal{A} \to \mathbb{C};\ A \mapsto \frac{1}{\epsilon^2} \sum_{\alpha,j} [\mathrm{tr}(\tau_j A(e_{x,\alpha}^\epsilon))]^2 \tag{I.2.5. 1}$$

where τ_j is a basis of $\mathrm{Lie}(G)$ with normalization $\mathrm{tr}(\tau_j \tau_k) = -N\delta_{jk}$ and $e_{x,\alpha}^\epsilon(t) = e_{x,\alpha}(\epsilon t)$. Using smoothness of A it easy to see that (I.2.3. 10) can be expanded in a convergent Taylor series with respect to ϵ with zeroth order component $\sum_{j,e_\alpha} |A_a^j(x) \dot{e}_{x,\alpha}^a(0)|^2$ whence $F_x^\epsilon \in \mathrm{Cyl}(\overline{X})$ separates our given $A \neq A'$. The proof for $\mathcal{A}$ replaced by $\mathcal{A}/\mathcal{G}$ is similar and was given by Giles [96] and will not be repeated here. In that proof it is important that G is compact.

We thus have the following abstract situation: A collection $\mathcal{C} = \mathrm{Cyl}(\overline{X})$ of bounded complex valued functions on a set $X = \mathcal{A}$ including the constants which separate the points of X. The set X maybe equipped with its own topology (e.g. the Sobolov topology that we defined in section I.1) but this will be irrelevant for the following result which is an abstract property of Abelean unital C^*-algebras.

Theorem I.2.11
Let $\mathcal{C}$ be a collection of real-valued, bounded functions on a set X which contain the constants and separate the points of X. Let $\overline{\mathcal{C}}$ be the Abelean, unital C^*- algebra generated from $\mathcal{C}$ by pointwise addition, multiplication, scalar multiplication and complex conjugation, completed in the sup-norm. Then the image of X under its natural embedding into the Gel'fand spectrum $\overline{X}$ of $\overline{\mathcal{C}}$ is dense with respect to the Gel'fand topology on the spectrum.

Proof of Theorem I.2.11:
Consider the following map

$$J:\ X \to \overline{X};\ x \mapsto J_x \text{ where } J_x(f) := f(x)\ \forall\ f \in \overline{\mathcal{C}} \tag{I.2.5. 2}$$

This is an injection since $J_x = J_{x'}$ implies in particular $f(x) = f(x')$ for all $f \in \mathcal{C}$, thus $x = x'$ since $\mathcal{C}$ separates the points of X by assumption, hence J provides an embedding.

Let $\overline{J(X)}$ be the closure of $J(X)$ in the Gel'fand topology on $\overline{X}$ of pointwise convergence on $\overline{\mathcal{C}}$. Suppose that $\overline{X} - \overline{J(X)} \neq \emptyset$ and take any $\chi \in \overline{X} - \overline{J(X)}$. Since $\overline{X}$ is a compact Hausdorff space we find $a \in C(\overline{X})$ such that $1 = a(\chi) \neq a(J_x) = 0$ for any $x \in X$ by Urysohn's lemma. (In Hausdorff spaces one point sets are closed, hence $\{\chi\}$ and $\overline{J(X)}$ are disjoint closed sets and finally compact Hausdorff spaces are normal spaces).

Since the Gel'fand map $\bigvee : \ \overline{\mathcal{C}} \to C(\overline{X})$ is an isometric isomorphism we find $f \in \overline{\mathcal{C}}$ such that $\check{f} = a$. Hence $0 = a(J_x) = \check{f}(J_x) = J_x(f) = f(x)$ for all $x \in X$, hence $f = 0$, thus $a \equiv 0$ contradicting $a(\chi) = 1$. Therfore χ in fact does not exist whence $\overline{X} = \overline{J(X)}$.

$\Box$

Of course in our case $\overline{\mathcal{C}} = \overline{\mathrm{Cyl}(\overline{\mathcal{A}})}$ and $\overline{X} = \overline{\mathcal{A}}$.

Our next result is actually much stronger than merely showing that $\mathcal{A}$ is contained in a measurable subset of $\overline{\mathcal{A}}$ of μ_0-measure zero.

Let e be an edge and if $e(t)$ is a representative curve then consider the family of segments e_s with $e_s(t) := e(st)$, $s \in [0,1]$. Consider the map

$$h^e : \ \overline{\mathcal{A}} \to \mathrm{Fun}([0,1], G); \ A \mapsto h^e_A \text{ where } h^e_A(s) := A(e_s) \tag{I.2.5. 3}$$

The set $\mathrm{Fun}([0,1], G)$ of all functions from the interval $[0,1]$ into G (no continuity assumptions) can be thought of as the uncountable direct product $G^{[0,1]} := \prod_{s\in[0,1]} G$ via the bijection $E : \ \mathrm{Fun}([0,1], G) \to G^{[0,1]}; \ h \to (h_s := h(s))_{s\in[0,1]}$. The latter space can be equipped with the Tychonov topology generated by the open sets on $G^{[0,1]}$ which are generated from the sets $P_s^{-1}(U_s) = [\prod_{s'\neq s} G] \times U_s$ (where $U_s \subset G$ is open in G) by finite intersections and arbitrary unions. Here $P_s : \ G^{[0,1]} \to G$ is the natural projection. Now the pre-image of such sets under h^e is given by

$$\begin{aligned} & (h^e)^{-1}(P_s^{-1}(U_s)) = \{A \in \overline{\mathcal{A}}; \ h^e_A \in p_s^{-1}(U_s)\} \\ = & \ \{A \in \overline{\mathcal{A}}; \ h^e_A(s) \in U_s, \ h^e_A(s') \in G \text{ for } s' \neq s\} \\ = & \ \{A \in \overline{\mathcal{A}}; \ A(e_s) \in U_s\} = p_{e_s}^{-1}(U_s) \end{aligned} \tag{I.2.5. 4}$$

where $p_{e_s} : \ \overline{\mathcal{A}} \to \mathrm{Hom}(e_s, G)$ is the natural projection in $\overline{\mathcal{A}}$. Since $\overline{\mathcal{A}}$ is equipped with the Tychonov topology, the maps p_{e_s} are continuous and since $\overline{\mathcal{A}}$ is equipped with the Borel σ algebra, continuous functions (pre-images of open sets are open) are automatically measurable (pre-images of open sets are measurable. Hence we have shown that h^e is a measurable map.

Let f be a function on $G^{[0,1]}$, that is, a complex valued function $h \mapsto f(\{h_s\}_{s\in[0,1]})$. We have an associated map of the form (I.2.1. 16), that is, $\rho_{l^e} : \ X_{l^e} \to G^{[0,1]}; \ A_{l^e} \mapsto (A_{l^e}(e_s) = h^e_A(s))_{s\in[0,1]}$ where l^e is the subgroupoid generated by the algebraically independent edges e_s. Thus $h^e = \rho_{l^e} \circ p_{l^e}$. The push-forward of the uniform measure $\nu := h^e_*\mu_0 = \mu_0 \circ (h^e)^{-1}$ is then the measure on $G^{[0,1]}$ given by

$$\begin{aligned} \int_{G^{[0,1]}} d\nu(h) f(h) &= \mu_0((h^e)^* f) = \mu_{0l^e}(\rho^*_{l^e} f) = \int_{G^{[0,1]}} \prod_{s\in[0,1]} d\mu_H(h_{e_s}) f(\{h_{e_s}\}_{s\in[0,1]}) \\ &\equiv \int_{G^{[0,1]}} \prod_{s\in[0,1]} d\mu_H(h_s) f(\{h_s\}_{s\in[0,1]}) \end{aligned} \tag{I.2.5. 5}$$

Theorem I.2.12
The measure μ_0 is supported on the subset D_e of $\overline{\mathcal{A}}$ defined as the set of those $A \in \overline{\mathcal{A}}$ such that h^e_A is nowhere continuous on $[0,1]$.

Proof of Theorem I.2.12:
Trivially

$$\begin{aligned} D_e &= \{A \in \overline{\mathcal{A}};\ h^e_A \text{ nowhere continuous in } [0,1]\} \\ &= (h^e)^{-1}(\{h \in G^{[0,1]};\ s \mapsto h_s \text{ nowhere continuous in } [0,1]\} =: (h^e)^{-1}(D) \end{aligned} \quad \text{(I.2.5. 6)}$$

If we can show that D contains a measurable set of $\nu-$measure one or that $G^{[0,1]} - D$ is contained in a measurable set D' of $\nu-$measure zero then we have shown that D_e contains a measurable set $D'_e = (h^e)^{-1}(G^{[0,1]} - D')$ of measure one because $\mu_0(D_e) = [\mu_0 \circ (h^e)^{-1}](G^{[0,1]} - D') = \nu(G^{[0,1]} - D') = 1$ and because h^e is measurable (since $G^{[0,1]}$ is equipped with the Borel $\sigma-$algebra). In other words, D_e will be a support for μ_0.

Let us then show that $G^{[0,1]} - D = \{h \in G^{[0,1]};\ \exists s_0 \in [0,1] \ni\ h \text{ continuous at } s_0\}$ is contained in a measurable set of $\nu-$measure zero. Let $h_0 \in G^{[0,1]} - D$, then we find $s_0 \in [0,1]$ such that h_0 is continuous at s_0. Fix any $0 < r < 1$ and consider an open cover of G by sets U with Haar measure $\mu_H(U) = r$. Since G is compact, we find a finite subcover, say $U_1, .., U_N$. Now there is $k_0 \in \{1, .., N\}$ such that $h_0(s_0) \in U_{k_0}$. By definition of continuity at a point we find an open interval $I \subset [0,1]$ such that $h(I) \subset U_{k_0}$. This motivates to consider the subsets $S_k := \{h \in G^{[0,1]};\ \exists I \subset [0,1] \text{ open } \ni h(I) \subset U_k\} \subset G^{[0,1]}$ and obviously $h_0 \in S_{k_0}$. Our aim is to show that these sets are contained in measure zero sets.

Let $B(q, 1/m) := \{s \in [0,1];\ |s-q| < 1/m\}$ with $q \in \mathbb{Q},\ m \in \mathbb{N}$. It is easy to show that these sets are a countable basis for the topology for $[0,1]$ (every open set can be obtained by arbitrary unions and finite intersections). Hence any open interval is given as a countable union of these open balls, i.e. $I = \bigcup_{B(q,m)\subset I} B(q,m)$. Since $h(I \cup J) = h(I) \cup h(J)$ we have

$$\begin{aligned} S_k &= \{h \in G^{[0,1]};\ \exists I \subset [0,1] \ni \bigcup_{B(q,m)\subset I} h(B(q,m)) \subset U_k\} = \bigcup_{(q,m)\in(\mathbb{Q}\times\mathbb{N})_k} S_{k,q,m} \\ S_{k,q,m} &:= \{h \in G^{[0,1]};\ h(B(q,m)) \subset U_k\} \end{aligned} \quad \text{(I.2.5. 7)}$$

where $(\mathbb{Q} \times \mathbb{N})_k$ are defined to be the subsets of rational and natural numbers (q,m) respectively such that $S_{U_k,q,m} \neq \emptyset$. (We could also remove that restriction).

We now show that $S_{k,q,m}$ is contained in a measure zero set. Let (s_n) be a sequence of points in $B(k,q,m)$. Then $S_{k,q,m} \subset \{h \in G^{[0,1]};\ h(s_n) \in U_k\ \forall s_n\} = \cap_n\{h \in G^{[0,1]};\ h(s_n) \in U_k\}$. Now the sets $\{h \in G^{[0,1]};\ h(s_n) \in U_k\} = P_s^{-1}(U_k)$ are measurable because P_s is continuous and U_k is open, hence so is $\cap_n\{h \in g^{[0,1]};\ h(s_n) \in U_k\}$. But

$$\nu(\cap_n\{h \in G^{[0,1]};\ h(s_n) \in U_k\}) = \nu([\prod_{s\neq s_n} G] \times [\prod_n U_k]) = \prod_n \mu_H(U_k) = \prod_n r = 0 \quad \text{(I.2.5. 8)}$$

since $r < 1$. Hence $S_{k,q,m}$ is contained in a measure zero subset and since ν is $\sigma-$additive also S_k is since (I.2.5. 7) is a countable union.

Finally, any $h_0 \in G^{[0,1]} - D$ is contained in one of the S_k, thus $G^{[0,1]} - D \subset \bigcup_{k=1}^N S_k$ is contained in a measurable subset of measure zero.
□

I.2.6 Spin – Network Functions, Loop Representation, Gauge – and Diffeomorphism Invariance of μ_0 and Ergodicity

In order to study the ergodicity properties of μ_0 we need to introduce an important concept, the so-called spin-network basis. We will distinguish between gauge variant and gauge invariant spin-network states. For representation theory on compact Lie groups, the Peter&Weyl theorem and Haar measures the reader is referred to [146].

Definition I.2.20
Fix once and for all a representative from each equivalence class of irreducible representations of the compact Lie group G and denote the collection of these representatives by Π. Let $l = l(\gamma)$ be given. Associate with every edge $e \in E(\gamma)$ a non-trivial, irreducible representation $\pi_e \in \Pi$ which we assemble in a vector $\vec{\pi} = (\pi_e)_{e\in E(\gamma)}$.
i)
The gauge variant spin-network functions are given by

$$T_{\gamma,\vec{\pi},\vec{m},\vec{n}}: \ \overline{\mathcal{A}} \to \mathbb{C}; \ A \mapsto \prod_{e\in E(\gamma)} \sqrt{d_{\pi_e}}[\pi_e(A(e))]_{m_e n_e} \tag{I.2.6. 1}$$

where d_π denotes the dimension of π and $\vec{m} = \{m_e\}_{e\in E(\gamma)}, \vec{n} = \{n_e\}_{e\in E(\gamma)}$ with $m_e, n_e = 1,.., d_{\pi_e}$ label the matrix elements of the representation.
ii)
Given a vertex $v \in V(\gamma)$ consider the subsets of edges given by $E^b_v(\gamma) := \{e \in E(\gamma);\ b(e) = v\}$ and $E^f_v(\gamma) := \{e \in E(\gamma);\ f(e) = v\}$. For each $v \in V(\gamma)$, consider the tensor product representation

$$(\otimes_{e\in E^b_v(\gamma)}\pi_e) \otimes (\otimes_{e\in E^f_v(\gamma)}\pi^c_e) \tag{I.2.6. 2}$$

where $h \mapsto \pi^c(h) := \pi(h^{-1})^T$ denotes the representation contragredient to π ($(.)^T$ denotes matrix transposition). Since G is compact, every representation is completely reducible and decomposes into an orthogonal sum of irreducible representations (not necessarily mutually inequivalent). Let $\mathcal{I}_v(\vec{\pi}, \pi'_v)$ be the set of all representations that appear in that decomposition of (I.2.6. 1) and which are equivalent to $\pi'_v \in \Pi$ with $\pi_t \in \Pi$ a representative of the trivial representation. An element $I_v \in \mathcal{I}_v(\vec{\pi}, \pi'_v)$ is called an intertwiner and we assemble a given choice of intertwiners into a vector $\vec{I} = (I_v)_{v\in V(\gamma)}$. By construction, we can project the representation (I.2.6. 2) into the representation $I_v \in \mathcal{I}_v(\vec{\pi}, \pi')$ by contracting (I.2.6. 2) with a corresponding intertwiner. Since the function

$$A \mapsto (\otimes_{e\in E^b_v(\gamma)}\pi_e(A(e))) \otimes (\otimes_{e\in E^f_v(\gamma)}\pi_e(A(e))) \tag{I.2.6. 3}$$

transforms in the representation (I.2.6. 2) under gauge transformations at v it therefore transforms in the representation I_v at v when contracted with the intertwiner $I_v \in \mathcal{I}_v(\vec{\pi}, \pi'_v)$. We now take the function

$$A \mapsto \otimes_{e\in E(\gamma)}\pi_e(A(e)) \tag{I.2.6. 4}$$

and for each vertex v consider the subproduct (I.2.6. 2) and then contract with an appropriate intertwiner I_v. The result is a cylindrical function on $\overline{\mathcal{A}}$ over $l = l(\gamma)$ which we denote by $T_{\gamma,\vec{\pi},\vec{I}}(A)$ and which transforms in the representation I_v at v. If we vary the π'_v, I_v then the set of functions $T_{\gamma,\vec{\pi},\vec{I}}$ span the same vector space as the space of functions $T_{\gamma,\vec{\pi},\vec{m},\vec{n}}$. In particular, we may take these functions to be normalized with respect to $\mathcal{H}_0$.

The gauge invariant spin network functions result when we restrict the π'_v to be trivial, that is, to equal π_t with the convention that $T_{\gamma,\vec{\pi},\vec{I}}$ vanishes if $\mathcal{I}_v(\vec{\pi}, \pi^t) = \emptyset$ for any $v \in V(\gamma)$. Since these functions are gauge invariant, we may consider them as functions $T_{\gamma,\vec{\pi},\vec{I}}: \ \overline{\mathcal{A}/\mathcal{G}} \to \mathbb{C}$.

Since spin-network functions are fundamental for what follows, let us discuss an example to make the definition clear:

Example
Consider the case of ultimate interest $G = SU(2)$ whose irreducible representations are labelled by non-negative, half-integral *spin* quantum numbers (from which the name "spin-network" originates). Consider a graph γ consisting of N edges $e_I,\ I = 1,..,N$ and two vertices v_1, v_2 such that $b(e_I) = v_1, f(e_I) = v_2$ for $1 \le I \le M$ and $b(e_I) = v_2, f(e_I) = v_1$ for $M+1 \le I \le N$ for some $1 < M < N$.

Associate with e_I an irreducible representation of $SU(2)$ labelled by the spin quantum number j_I. Under gauge transformations the function

$$A \mapsto \prod_{I=1}^{N} \pi_{j_I}(A(e_I))_{m_I n_I} = [\otimes_{I=1}^N \pi_{j_I}(A(e_I))]_{m_1,..,m_N;n_1,..,n_N} \qquad \text{(I.2.6. 5)}$$

with $m_I, n_I = 1,..,2j_I+1 = d_{\pi_{j_I}}$ is mapped into

$$\begin{aligned}
A &\mapsto [\otimes_{I=1}^N \pi_{j_I}(g(b(e_I))A(e_I)g(f(e_I))^{-1})]_{m_1,..,m_N;n_1,..,n_N} \qquad \text{(I.2.6. 6)}\\
&= [\{\otimes_{I=1}^M \pi_{J_I}(g(v_1))\pi_{j_I}(A(e_I))\pi_{j_I}(g(v_2))^{-1}\} \otimes \\
&\quad \otimes\{\otimes_{I=M+1}^N \pi_{J_I}(g(v_2))\pi_{j_I}(A(e_I))\pi_{j_I}(g(v_1))^{-1}\}]_{m_1,..,m_N;n_1,..,n_N}\\
&= [\{\otimes_{I=1}^M \pi_{J_I}(g(v_1))\} \otimes \{\otimes_{I=M+1}^N \pi_{J_I}(g(v_1))^{-1}\}]_{m_1,..,m_M,l_{M+1},..,l_N;k_1,..,k_M,n_{M+1},..,n_N} \times\\
&\quad \times[\otimes_{I=1}^N \pi_{j_I}(A(e_I))]_{k_1,..,k_N;l_1,..,l_N} \times\\
&\quad \times[\{\otimes_{I=1}^M \pi_{J_I}(g(v_2))^{-1}\} \otimes \{\otimes_{I=M+1}^N \pi_{J_I}(g(v_2))\}]_{l_1,..,l_M,m_{M+1},..,m_N;n_1,..,n_M,k_{M+1},..,k_N}\\
&= [\{\otimes_{I=1}^M \pi_{J_I}(g(v_1))\} \otimes \{\otimes_{I=M+1}^N \pi^c_{J_I}(g(v_1))\}]_{m_1,..,m_M,n_{M+1},..,n_N;k_1,..,k_M,l_{M+1},..,l_N} \times\\
&\quad \times[\otimes_{I=1}^N \pi_{j_I}(A(e_I))]_{k_1,..,k_N;l_1,..,l_N} \times\\
&\quad \times[\{\otimes_{I=1}^M \pi^c_{J_I}(g(v_2))\} \otimes \{\otimes_{I=M+1}^N \pi_{J_I}(g(v_2))\}]_{n_1,..,n_M,m_{M+1},..,m_N;l_1,..,l_M,k_{M+1},..,k_N}
\end{aligned}$$

from which we see how the contragredient representation enters the stage. Now $SU(2)$ is special in the sense that a representation and its contragredient one are equivalent which follows from $g^c = \tau_2 g \tau_2^{-1}$ for any $g \in SU(2)$ where $-\tau_2 = i\sigma_2$ is the spinor metric. We are thus lead to consider tensor products of the form

$$\begin{aligned}
&[\{\otimes_{I=1}^M \pi_{J_I}(g)\} \otimes \{\otimes_{I=M+1}^N \pi^c_{j_I}(g)\}] = [\{\otimes_{I=1}^M \pi_{J_I}(1)\} \otimes \{\otimes_{I=M+1}^N \pi_{j_I}(\tau_2)\}] \cdot\\
\cdot\ &[\otimes_{I=1}^N \pi_{J_I}(g)] \cdot [\{\otimes_{I=1}^M \pi_{J_I}(1)\} \otimes \{\otimes_{I=M+1}^N \pi_{j_I}(\tau_2)^{-1}\} \qquad \text{(I.2.6. 7)}
\end{aligned}$$

In order to decompose the tensor product representation $j_1 \otimes j_2 \otimes .. \otimes j_N$ into irreducibles we must agree on a *recoupling scheme*, that is, we must decide on a bracketing of this tensor product. We choose $(..((j_1 \otimes j_2) \otimes j_3) \otimes ..) \otimes j_N$ and apply the Clebsch-Gordan theorem $j_1 \otimes j_2 = j_1 + j_2 \oplus j_1 + j_2 - 1 \oplus .. \oplus |j_1 - j_2|$ starting from the inner most bracket and working our way outwards. For instance in the case $N = 3$ we have

$$\begin{aligned}
&(j_1 \otimes j_2) \otimes j_3 = (j_1 + j_2 \oplus .. \oplus |j_1 - j_2|) \otimes j_3\\
= \ &(j_1 + j_2 + j_3 \oplus .. \oplus |j_1 + j_2 - j_3|) \oplus (j_1 + j_2 - 1 + j_3 \oplus .. \oplus |j_1 + j_2 - 1 - j_3|) \oplus ..(|j_1 - j_2| + j_3 \oplus .. \oplus ||j_1 -
\end{aligned}$$

Notice that all appearing representations, even those that appear with multiplicity higher than one (and are therefore mutually equivalent), are realized on mutually orthogonal subspaces of the

$(2j_1+1)(2j_2+1)(2j_3+1)$ dimensional representation space of $j_1 \otimes j_2 \otimes j_3$. We see that in this case $N=3$ $\mathcal{I}_{v_{1,2}}(\vec{j}, j=0)$ is empty unless $j_3 \in \{j_1+j_2, .., |j_1-j_2|\}$ and if that is the case there is only one trivial representation contained in (I.2.6. 8) no matter how large j_1, j_2 are. If $N>3$ this is no longer true, the space of spin-network states on graphs with at least one vertex of valence larger than three is generically more than one-dimensional for given values of the j_I.

We see that the theory of spin-network states for $SU(2)$ is largely governed by the representation theory of $SU(2)$ and Clebsh-Gordan coefficients which give the precise numerical coefficients in the orthogonal sums (I.2.6. 8). In particular, changing of recoupling schemes gives rise to the complicated $3Nj$ symbols which can be decomposed into $6j$ symbols. It is this complicated recoupling theory that determines the spectrum of the volume operator, see [147] for an introduction using the terminology of the present review.

Once we have then isolated all possible trivial representations in the decomposition (I.2.6. 7) we insert one of them back into (I.2.6. 6) in place of (I.2.6. 7) and have found a suitable, gauge invariant intertwiner. This we do for all possible (mutually orthogonal) intertwiners and vertices and have then found all possible gauge invariant states over the given γ given the j_I. This concludes our example.

The importance of spin-network functions is that they provide a basis for $\mathcal{H}^0$.

Theorem I.2.13

i)

The gauge variant spin-network states provide an orthonormal basis for the Hilbert space $L_2(\overline{\mathcal{A}}, d\mu_0)$.

ii)

The gauge invariant spin-network states provide an orthonormal basis for the Hilbert space $L_2(\overline{\mathcal{A}/\mathcal{G}}, d\mu_0)$.

Proof of Theorem I.2.13:

i)

The inner product on $L_2(\overline{\mathcal{A}}, d\mu_0)$ is defined by

$$< f, f' >_{L_2(\overline{\mathcal{A}}, d\mu_0)} := \Lambda_{\mu_0}(\overline{f} f') \tag{I.2.6. 9}$$

where Λ_{μ_0} is the positive linear functional on $C(\overline{\mathcal{A}})$ determined by μ_0 via the Riesz representation theorem. The cylinder functions of the form $p_l^* f_l,\ f_l \in C(X_l)$ are dense in $C(\overline{\mathcal{A}})$ (in the sup-norm) and since $\overline{\mathcal{A}}$ is a (locally) compact Hausdorff space and μ_0 comes from a positive linear functional on the space of continuous functions on $\overline{\mathcal{A}}$ (of compact support), these functions are dense in $L_2(\overline{\mathcal{A}}, d\mu_0)$ (in the L_2 norm $||f||_2 =< f, f >^{1/2}$, see e.g. [145]). It follows that $L_2(\overline{\mathcal{A}}, d\mu_0)$ is the completion of $\mathrm{Cyl}(\overline{\mathcal{A}})$ in the L_2 norm. Now

$$\mathrm{Cyl}(\overline{\mathcal{A}}) = \bigcup_{l \in \mathcal{L}} p_l^* C(X_l) \tag{I.2.6. 10}$$

and since by the same remark $C(X_l)$ is dense in $L_2(X_l, d\mu_{0l})$ it follows that

$$L_2(\overline{\mathcal{A}}, d\mu_0) = \overline{\bigcup_{l \in \mathcal{L}} p_l^* L_2(X_l, d\mu_{0l})} \tag{I.2.6. 11}$$

Now by definition $(\rho_l)_* \mu_{0l} = \otimes_{e \in E(\gamma)} \mu_H$ for $l = l(\gamma)$ so that $L_2(X_l, d\mu_{0l})$ is isometric isomorphic with $L_2(G^{|E(\gamma)|}, \otimes^{|E(\gamma)|} d\mu_H)$ which in turn is isometric isomorphic with $\otimes_{e \in E(\gamma)} L_2(G, d\mu_H)$ since $\otimes^{|E(\gamma)|}\mu_H$ is a finite product of measures. By the Peter&Weyl theorem the matrix element functions

$$\pi_{mn}:\ G \to \mathbb{C};\ h \mapsto \sqrt{d_\pi}\pi_{mn}(h), \pi \in \Pi,\ m, n = 1, .., d_\pi \tag{I.2.6. 12}$$

form an orthonormal basis of $L_2(G, d\mu_H)$ for any compact gauge group G, that is,

$$< \pi_{mn}, \pi'_{m'n'} >:= \int_G d\mu_H(h)\overline{\pi_{mn}(h)}\pi'_{m'n'}(h) = \frac{\delta_{\pi\pi'}\delta_{mm'}\delta_{nn'}}{d_\pi} \tag{I.2.6. 13}$$

This shows that functions of the form (I.2.6. 1) span $L'_2(X_l, d\mu_{0l}) :\cong \otimes^{|E(\gamma)|}L'_2(G, d\mu_H)$ where $L'_2(G, d\mu_H)$ is the closed linear span of the functions π_{mn} with $\pi \neq \pi_t$ (only non-trivial representations allowed).

It remains to prove 1) that $p_l^* L'_2(X_l, d\mu_{0l}) \perp p_{l'}^* L'_2(X_{l'}, d\mu_{0l'})$ unless $l = l'$ and 2) that $L_2(X_l, d\mu_{0l}) = \overline{\oplus_{l' \prec l} L'_2(X_{l'}, d\mu_{0l'})}$. where completion is with respect to $L_2(X_l, d\mu_{0l})$.
To see the former, notice that if $l = l(\gamma) \neq l' = l(\gamma')$ there is $l, l' \prec l'' := l(\gamma \cup \gamma')$. Since $\gamma \neq \gamma'$ are piecewise analytic, there must be an edge $e \in E(\gamma)$ which contains a segment $s \subset e$ which is disjoint from γ' (reverse the roles of γ, γ' if necessary) and this segment is certainly contained in $\gamma \cup \gamma'$. Let $f_l \in L'_2(X_l, d\mu_{0l}), f_{l'} \in L'_2(X_{'l}, d\mu_{0l'})$ then

$$< p_l^* f_l, p_{l'}^* f_{l'} >= \mu_{0l''}(\overline{p_{l''l}^* f_l} p_{l''l'}^* f_{l'}) = 0 \tag{I.2.6. 14}$$

since $p_{l''l}^* f_l, p_{l''l'}^* f_{l'}$ are (Cauchy sequences of) functions of the form (I.2.6. 1) over $\gamma \cup \gamma'$ where the dendence on s of the former function is through a non-trivial representation and of the latter through a trivial representation, so the claim follows from formula (I.2.6. 13).
To see the former, observe that $L_2(G, d\mu_H) = \overline{L'_2(G, d\mu_H) \oplus \text{span}(\{1\})}$ and that a function cylindrical over γ which depends on $e \in E(\gamma)$ through the trivial representation is cylindrical over $\gamma - e$ as well.

Summarizing, if we define $\mathcal{H}_l^0 := p_l^* L_2(X_l, d\mu_{0l}),\ \mathcal{H}^{0l} := p_l^* L'_2(X_l, d\mu_{0l})$ then

$$\mathcal{H}^0 = \overline{\bigcup_{l\in\mathcal{L}} p_l^* \mathcal{H}_l^0} = \overline{\oplus_{l\in\mathcal{L}} p_l^* \mathcal{H}^{0l}} \tag{I.2.6. 15}$$

ii)
The assertion follows easily from i) and the fact that $L_2(\overline{\mathcal{A}/\mathcal{G}}, d\mu_0)$ is simply the restriction of $L_2(\overline{\mathcal{A}}, d\mu_0)$ to the gauge invariant subspace: That subspace is the closed linear span of gauge invariant spin-network states by i) and the specific choice that we have made in definition I.2.20 shows that they form an orthonormal system since we have chosen them to be normalized and the intertwiners to be projections onto mutually orthogonal subspaces of a tensor product representation space of G. More specifically, the inner product between two spin network functions $T_{\gamma,\vec{\pi},\vec{I}}, T_{\gamma',\vec{\pi}',\vec{I}'}$ is nonvanishing only if $\gamma = \gamma'$ and $\vec{\pi} = \vec{\pi}'$. In that case, consider $v \in V(\gamma)$ and assume w.l.g. that all edges $e_1, .., e_N$ incident at v are outgoing. An intertwiner $I_v \in \mathcal{I}_v(\vec{\pi}, \pi_t)$ can be thought of as a vector $I_v^{n_1,..,n_N} := (I_v)_{m_1^0,..,m_N^0;n_1,..,n_N}$ in the representation space of the representation $\otimes_{I=1}^N \pi_I$ where m_I^0 are some matrix elements that we fix once and for all. Since I_v is a trivial representation and in particular represents $1_G = (1_G)^T$ we have $(I_v)_{m_1^0,..,m_N^0;n_1,..,n_N} = I_v^{n_1,..,n_N} := (I_v)_{n_1,..,n_N;m_1^0,..,m_N^0}$, moreover the intertwiners are real valued because the functions $\pi_{mn}(h)$ depend analytically on h and 1_G is real valued. Now the spin-network state restricted to its dependence on $e_1, .., e_N$ is of the form

$$I_v^{n_1,..,n_N}[\otimes_{I=1}^N \pi_I(A(e_I))]_{n_1,..,n_N;k_1,..,k_N} \tag{I.2.6. 16}$$

It follows from (I.2.6. 13) that the inner product between $T_{\gamma,\vec{\pi},\vec{I}}, T_{\gamma,\vec{pi},\vec{I}'}$ will be proportional to

$$I_v^{n_1,..,n_N}(I')_v^{n_1,..,n_N} = [(I_v)(I'_v)]_{m_1^0,..,m_N^0;m_1'^0,..,m_N'^0} \propto \delta_{I_v I'_v} \tag{I.2.6. 17}$$

(if $I_v = I'_v$ then $m_I^0 = m_I^{0\prime}$ by construction) since the I_v are representations on mutually orthogonal subspaces.

$\square$

We remark that the spin-network basis is not countable because the set of graphs in σ is not countable, whence $\mathcal{H}^0$ is not separable. We will see that this is even the case after moding out by spatial diffeomorphisms although one can argue that after moding out by diffeomorphisms the remaining space is an orthogonal, uncountably infinite sum of superselected, mutually isomorphic, separable Hilbert spaces [119].

Definition I.2.21

The gauge variant spin-network representation is a vector space $\tilde{\mathcal{H}}^0$ of complex valued functions

$$\psi : \ \mathcal{S} \to \mathbb{C}; \ s \mapsto \psi(s) \tag{I.2.6. 18}$$

where $\mathcal{S}$ is the set of quadruples $(\gamma, \vec{\pi}, \vec{m}, \vec{n})$ which label a spin-network state. Likewise, the loop representation is the gauge invariant spin-network representation defined analogously. This vector space is equipped with the scalar product

$$< \psi, \psi' >_{\tilde{\mathcal{H}}^0} := \sum_{s \in \mathcal{S}} \overline{\psi(s)} \psi'(s) \tag{I.2.6. 19}$$

between square summable functions.

Clearly the uncountably infinite sum (I.2.6. 19) converges if and only if $\psi(s) = 0$ except for countably many $s \in \mathcal{S}$. The next corollary shows that the connection representation that we have been dealing with so far and the spin-network representation are in a precise sense Fourier transforms of each other where the role of the kernel of the transform is played by the spin-network functions.

Corollary I.2.2 *The spin-network (or loop) transform*

$$T : \ \mathcal{H}^0 \to \tilde{\mathcal{H}}^0; \ f \mapsto \tilde{f}(s) := < T_s, f >_{\mathcal{H}^0} \tag{I.2.6. 20}$$

is a unitary transformation between Hilbert spaces with inverse

$$(T^{-1}\psi)(A) := \sum_{s \in \mathcal{S}} \psi(s) T_s(A) \tag{I.2.6. 21}$$

Proof of Corollary I.2.2:
If $f \in \mathcal{H}^0$ then

$$f = \sum_{s \in \mathcal{S}} < T_s, f > T_s \tag{I.2.6. 22}$$

since the T_s form an orthonormal basis (Bessel's inequality is saturated). Since the T_s form an orthonormal system we conclude that $||f||^2 = \sum_s | < T_s, f > |^2$ converges, meaning in particular that $< T_s, f >= 0$ except for finitely many $s \in \mathcal{S}$. It follows that $||Tf||^2 := \sum_s |\tilde{f}(s)|^2 = ||f||^2$ which shows that T is a partial isometry. Comparing (I.2.1. 29) and (I.2.1. 30) we see that $T^{-1}\tilde{f} = f$ is indeed the inverse of T. Finally again by the orthogonality of the T_s we have $||T^{-1}\psi||^2 = \sum_s |\psi(s)|^2 = ||\psi||^2$ so that T^{-1} a partial isometry as well. Since T is a bijection, T is actually an isometry. Notice that $\tilde{T}_s(s') = \delta_{s,s'}$.

$\square$

Whenever it is convenient we may therefore think of states either in the loop or the connection representation. In this review we will work entirely in the connection representation.

In the previous section we have investigated the topological and measure theoretical relation between $\mathcal{A}$ and $\overline{\mathcal{A}}$. In this section we will investigate the action of the gauge and diffeomorphism group on $\overline{\mathcal{A}}$. The uniform measure has two important further properties: it is invariant under both the gauge group $\overline{\mathcal{G}}$ and the Diffeomorphism group $\mathrm{Diff}^\omega(\sigma)$ (analytic diffeomorphisms). To see this, recall the action of $\overline{\mathcal{G}}$ on $\overline{\mathcal{A}}$ defined through its action on the subspaces X_l by $x_l \mapsto \lambda_g(x_l)$ with $[\lambda_g(x_l)](p) = g(b(p))x_l(p)g(f(p))^{-1}$ for any $p \in l$. This action has the feature to leave the X_l invariant for any $l \in \mathcal{L}$ and therefore lifts to $\overline{X}$ as $x \mapsto \lambda_g(x)$ with $[\lambda_g(x)](p) = g(b(p))x(p)g(f(p))^{-1}$ for any $p \in \mathcal{L}$. Likewise we have an action of $\mathrm{Diff}^\omega(\sigma)$ on $\overline{X}$ defined by

$$\delta^l:\ \mathrm{Diff}^\omega(\sigma) \times X_l \to X_{\varphi^{-1}(l)};\ (\varphi, x_l) \mapsto \delta^l_\varphi(x_l) = x_{\varphi^{-1}(l)} \tag{I.2.6. 23}$$

where $\varphi^{-1} = l(\varphi^{-1}(\gamma))$ if $l = l(\gamma)$. This action does not preserve the various X_l. The action on all of $\overline{X}$ is then evidently defined by

$$\delta:\ \mathrm{Diff}^\omega(\sigma) \times \overline{X} \to \overline{X};\ (\varphi, x = (x_l)_{l\in\mathcal{L}}) \mapsto \delta_\varphi(x) = (\delta^l_\varphi(x_l))_{l\in\mathcal{L}} \tag{I.2.6. 24}$$

Clearly $\delta_\varphi(x)$ is still an element of the projective limit since it just permutes the various x_l among each other. Moreover, $l \prec l'$ iff $\varphi^{-1}(l) \prec \varphi^{-1}(l')$ so the diffeomorphisms preserve the partial order on the label set. Therefore

$$p_{\varphi^{-1}(l')\varphi^{-1}(l)}(\delta^{l'}_\varphi(x_{l'}) = x_{\varphi^{-1}(l)} = \delta^l_\varphi(p_{l'l}(x_{l'}) \tag{I.2.6. 25}$$

for any $l \prec l'$, so we have equivariance

$$p_{\varphi^{-1}(l')\varphi^{-1}(l)} \circ \delta^{l'}_\varphi = \delta^l_\varphi \circ p_{l'l} \tag{I.2.6. 26}$$

It is now easy to see that for the push-forward measures we have $(\lambda_g)_*\mu_0 = \mu_0, (\delta_\varphi)_*\mu_0 = \mu_0$. For any $f = p_l^* f_l \in C(\overline{X}),\ f_l = \rho_l^* F_l \in C(X_l),\ F_l \in C(G^{|E(\gamma)|}),\ l = l(\gamma) \in \mathcal{L}$ we have

$$\begin{aligned}
\mu_0(\lambda_g^* f) &= \mu_0(p_l^*(\lambda_g^l)^* f_l) = \mu_{0l}((\lambda_g^l)^* f_l) \\
&= \int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(h_e)] F_l(\{g(b(e))h_e g(f(e))^{-1}\}_{e\in E(\gamma)}) \\
&= \int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(g(b(e))^{-1}h_e g(f(e)))] F_l(\{h_e\}_{e\in E(\gamma)}) \\
&= \int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(h_e)] F_l(\{h_e\}_{e\in E(\gamma)}) = \mu_0(f)
\end{aligned} \tag{I.2.6. 27}$$

where we have made a change of integration variables $h_e \to g(b(e))h_e g(f(e))^{-1}$ and used that the associated Jacobian equals unity for the Haar measure (translation invariance). Next

$$\begin{aligned}
\mu_0(\delta_\varphi^* f) &= \mu_0(p^*_{\varphi^{-1}(l)}(\delta^l_g)^* f_l) = \mu_{0\varphi^{-1}(l)}((\delta^l_\varphi)^* f_l) \\
&= \int_{G^{|E(\varphi^{-1}(\gamma))|}} [\prod_{e\in E(\varphi^{-1}(\gamma))} d\mu_H(h_e)] F_l(\{h_e\}_{e\in E(\varphi^{-1}(\gamma))}) \\
&= \int_{G^{|E(\gamma)|}} [\prod_{e\in E(\gamma)} d\mu_H(h_e)] F_l(\{h_e\}_{e\in E(\gamma)}) = \mu_0(f)
\end{aligned} \tag{I.2.6. 28}$$

where we have written $\{h_e\}_{e\in E(\varphi^{-1}(\gamma))} = \{h_{\varphi^{-1}(e)}\}_{e\in E(\gamma)}$ and have performed a simple relabelling $h_{\varphi^{-1}(e)} \to h_e$. It is important to notice that in contrast to other measures on some space of connections the "volume of the gauge group is finite": The space $C(\overline{\mathcal{A}/\mathcal{G}})$ is a subspace of $C(\overline{\mathcal{A}})$ and we may integrate them with the measure μ_0 which is the same as integrating them with the restricted measure. We do not have to fix a gauge and never have to deal with the problem of Gribov copies.

One may ask now why one does not repeat with the diffeomorphism group what has been done with the gauge group: Passing from analytic diffeomorphisms $\text{Diff}^\omega(\sigma)$ to distributional ones $\overline{\text{Diff}(\sigma)}$ and passing to the quotient space $(\overline{\mathcal{A}/\mathcal{G}})/\overline{\text{Diff}(\sigma)}$. There are two problems: First, in the case of $\overline{\mathcal{G}}$ there was a natural candidate for the extension $\mathcal{G} \to \overline{\mathcal{G}}$ but this is not the case for diffeomorphisms because distributional diffeomorphisms will not lie in any differentiability category any more and therefore are not diffeomorphisms in the strict sense. Secondly, as we will now show, even the analytic diffeomorphisms act ergodically on the measure space which means that there are no non-trivial invariant functions. Thus, one either has to proceed differently (e.g. downsizing rather than extending the diffeomorphism group), change the measure or solve the diffeomorphism constraint differently. We will select the third option in section I.3. It should be pointed out, however, that the last word of how to deal with diffeomorphism invariance has not been spoken yet. In a sense, it is one of the *key questions* for the following reason: The concept of a smooth spacetime should not have any meaning in a quantum theory of the gravitational field where probing distances beyond the Planck length must result in black hole creation which then evaporate in Planck time, that is, spacetime should be fundamentally discrete. But clearly smooth diffeomorphisms have no room in such a discrete quantum spacetime. The fundamental symmetry is probably something else, maybe a combinatorial one, that looks like a diffeomorphism group at large scales. Also, if one wants to allow for topology change in quantum gravity then talking about the diffeomorphism group for a fixed σ does not make much sense. We see that there is a tension between classical diffeomorphism invariance and the discrete structure of quantum spacetime which in our opinion has not been satisfactorily resolved yet and which we consider as one of the most important conceptual problems left open so far.

Let us then move on to establish ergodicity:
The above discussion reveals that as far as $\overline{\mathcal{G}}$ and $\text{Diff}(\sigma)$ are concerned we have the following abstract situation (see section III.5): We have a measure space with a measure preserving group action of both groups (so that the pull-back maps $\lambda_g^*, \delta_\varphi^*$ provide unitary actions on the Hilbert space) and the question is whether that action is ergodic. That is certainly not the case with respect to $\overline{\mathcal{G}}$ since the subspace of gauge invariant functions is by far not the span of the constant functions as we have shown.

Theorem I.2.14
The group $\mathit{Diff}_0^\omega(\sigma)$ of analytic diffeomorphisms on an analytic manifold σ connected to the identity acts ergodically on the measure space $\overline{\mathcal{A}}$ with respect to the Borel measure μ_0.

Proof of Theorem I.2.14:
The diffeomorphism group acts unitarily on $\mathcal{H}^0$ via

$$[\hat{U}(\varphi)f](A) = f(\delta_\varphi(A)) \tag{I.2.6. 29}$$

which means for spin-network states that $\hat{U}(\varphi)T_s = T_{\varphi^{-1}(s)}$ where

$$\varphi^{-1}(s) = (\varphi^{-1}(\gamma),\ \{\pi_{\varphi^{-1}(e)} = \pi_e\}_{e\in E(\gamma)},\ \{m_{\varphi^{-1}(e)} = m_e\}_{e\in E(\gamma)},\ \{n_{\varphi^{-1}(e)} = n_e\}_{e\in E(\gamma)} \tag{I.2.6. 30}$$

for $s = (\gamma, \vec{\pi}, \vec{m}, \vec{n})$. Let now $f = \sum_{s\in\mathcal{S}} c_s\ T_s \in \mathcal{H}^0$ be given with $c_s = 0$ except for countably many. Suppose that $\hat{U}(\varphi)f = f\ \mu_0-$a.e. for any $\varphi \in \text{Diff}_0^\omega(\sigma)$. Since $\mathcal{S}$ is left invariant by diffeomorphisms, this means that

$$\sum_s c_s T_{\varphi^{-1}(s)} = \sum_s c_{\varphi(s)} T_s = \sum_s c_s T_s \tag{I.2.6. 31}$$

for all φ. Since the T_s are mutually orthogonal we conclude that $c_s = c_{\varphi(s)}$ for all $\varphi \in \text{Diff}_0^\omega(\sigma)$. Now for any $s \neq s_0 = (\emptyset, \vec{0}, \vec{0}, \vec{0})$ the orbit $[s] = \{\varphi(s);\ \varphi \in \text{Diff}_0^\omega(\sigma)\}$ contains infinitely many different elements (take any vector field that does not vanish in an open set which contains the graph determined by s and consider the one parameter subgroup of diffeomorphisms determined by its integral curve – this is where we can make the restriction to the identity component). Therefore c_s =const. for infinitely many s. Since f is normalizable, this is only possible if const.= 0, hence $f = c_{s_0} T_{s_0}$ is constant μ_0-a.e. and therefore δ ergodic.
$\Box$
We see that the theorem would still hold if we would replace $\text{Diff}_0^\omega(\sigma)$ by any infinite subgroup D with respect to which each orbit $[s]$, $s \neq s_0$ is infinite. An example would be the case $\sigma = \mathbb{R}^D$ and D a discrete subgroup of the translation group given by integer multiples of translations by a fixed non-zero vector.

The theorem shows that the only vectors in $\mathcal{H}^0$ invariant under diffeomorphisms are the constant functions, hence we cannot just pass to that trivial subspace in order to solve the diffeomorphism constraint. The solution to the problem lies in passing to a larger space of functions, distributions over a subspace of $\mathcal{H}^0$ in which one can solve the constraint. The proof of the theorem shows already how that distributional space must look like: it must allow for uncountably infinite linear combinations of the form $\sum_s c_s T_s$ where c_s is a generalized knot invariant (i.e. $c_s = c_{\varphi s}$ for any φ, generalized because $\gamma(s)$ has in general self-intersections and is not a regular knot). This brings us to the next section.

I.3 Quantum Kinematics

This section is concerned with the following issues: In the previous section we have introduced a distributional extension $\overline{\mathcal{A}}$ of the space of smooth connections $\mathcal{A}$ which we choose as our quantum configuration space. We equipped it with a topology and the natural Borel $\sigma-$algebra that comes with it and have defined a natural measure μ_0 thereon. The measure is natural because it is invariant under both gauge transformations and spatial diffeomorphisms. However, in order to be physically meaningful we must show that the corresponding L_2 Hilbert space implements the correct adjointness relations and canonical commutation relations. This will be our first task. Next we must solve the quantum constraints by the methods of refined algebraic quantization (RAQ) to which we will give a brief introduction in section III.7. In order to do this we must define the constraints as closed, densely defined operators on the Hilbert space $\mathcal{H}^0$ and look for solutions in the algebraic dual of a certain subspace thereof. Since the solutions to the constraints are not elements of $\mathcal{H}^0$ as we already saw at the end of the previous section, one must define a new inner product on the space of solutions. We will do this in this section restricted to the kinematical constraints, that is, the Gauss and Diffeomorphism constraint. The inner product on the space of solutions of the Diffeomorphism constraint that we will derive by using RAQ methods is, however, only of mathematical interest because it is not possible to solve the Diffeomorphism and Hamiltonian constraint in two separate steps, the (dual) Hamiltonian constraint does not leave the space of diffeomorphism invariant distributions invariant.

I.3.1 Canonical Commutation – and Adjointness Relations

In this section as well as the two following ones it will not be important that $G = SU(2)$ or that σ is threedimensional, hence we will leave the discussion at the level of general compact, connected gauge groups G and $D-$dimensional analytic manifolds.

I.3.1.1 Classical Lie Algebra of Functions and Vector Fields on $\mathcal{A}$: Electric Fluxes

In order to convince ourselves that $\mathcal{H}^0 = L_2(\overline{\mathcal{A}}, d\mu_0)$ implements the correct adjointness and canonical commutation relations we must first decide on an appropriate set of classical functions that separate the points of the classical phase space $\mathcal{M}$. For the configuration space we have already seen several times that the holonomy functions $p \mapsto A(p) := h_p(A)$ with $p \in \mathcal{P}$ separate the points of the space of classical connections $\mathcal{A}$. We now have to look for appropriate momentum space functions.

Let S be an analytic, orientable, connected, embedded $(D-1)$-dimensional submanifold of σ (a surface) which we choose to be open. Since E^a_j is a vector density of weight one, the function $(*E_j)_{a_1..a_{D-1}} := E^c_j \epsilon_{ca_1..a_{D-1}}$ is a $(D-1)$-form which we may integrate in a background independent way over S, that is,

$$E_j(S) := \int_S *E_j \tag{I.3.1. 1}$$

These functions, which we will refer to as *electric flux* variables, certainly separate the space $\mathcal{E}$ of smooth electric fields on σ: To see this consider a surface of the form $S : (-1/2, 1/2)^{D-1} \to \sigma;\ (u_1, .., u_{D-1}) \mapsto S(u_1, .., u_{D-1})$ with analytic functions $S(u_1, .., u_{D-1})$ and let $S_\epsilon(u_1, .., u_{D-1}) := S(\epsilon u_1, .., \epsilon u_{D-1})$. Then (I.3.1. 1) becomes

$$\begin{aligned} E_j(S_\epsilon) &= \int_{(-\epsilon/2,\epsilon/2)^{D-1}} du_1..du_{D-1}\epsilon_{aa_1..a_{D-1}}(\partial S^{a_1}/\partial u_1)(u_1, .., u_{D-1}).. \\ &\quad ..(\partial S^{a_{D-1}}/\partial u_{D-1})(u_1, .., u_{D-1})E^a_j(S(u_1, .., u_{D-1})) \\ &= \epsilon^{D-1}\epsilon_{aa_1..a_{D-1}}(\partial S^{a_1}/\partial u_1)(0, .., 0)..(\partial S^{a_{D-1}}/\partial u_{D-1})(0, .., 0)E^a_j(S(0, .., 0)) + O(\epsilon^D) \end{aligned} \tag{I.3.1. 2}$$

where we have written the lowest order term in the Taylor expansion in the second line. It follows that

$$\lim_{\epsilon\to 0}\frac{E_j(S_\epsilon)}{\epsilon^{D-1}}=\epsilon_{aa_1..a_{D-1}}(\partial S^{a_1}/\partial u_1)(0,..,0)..(\partial S^{a_{D-1}}/\partial u_{D-1})(0,..,0)E^a_j(S(0,..,0))+O(\epsilon^D) \quad \text{(I.3.1. 3)}$$

and by varying S we may recover every component of $E^a_j(x)$ at $x=S(0,..,0)$.

The proposal then is to start classically from the set functions $A(p),E(S)$. Notice that in contrast to the holonomy functions the functions $E_j(S)$ do not have a simple behaviour under gauge transformations. This is not troublesome at the level of gauge-variant functions but will become a problem when passing to the gauge invariant sector. Fortunately, as we will see, one can construct gauge invariant functions from the $E_j(S)$ by certain limiting procedures and the amazing fact is that the corresponding operators continue to be well-defined in quantum theory. Thus, we will use the functions $E_j(S)$ directly only in an intermediate step in order to verify that the reality conditions and canonical brackets are correctly implemented.

I.3.1.2 Regularization of the Magnetic and Electric Flux Poisson Algebra

The reality conditions are simply that $A(p)$ is $G-$valued and that $E(S)=E_j(S)\tau_j$ is Lie$(G)-$valued, i.e., $E_j(S)$ is real valued. The Poisson brackets among $A(p),E(S)$ are, however, a priori ill-defined because the Poisson brackets that we derived in section I.1 required that the fields A,E be smeared in D directions by smooth functions while the functions $A(p),E(S)$ represent only one – and $(D-1)-$ dimensional smearings only. Therfore it is not possible to simply compute their Poisson brackets: The aim to have a background independent formulation of the quantum theory forces us to consider such singular smearings and prevents us from using the Poisson brackets on $\mathcal{M}$ directly. The strategy will therfore be to regularize the functions $A(p),E(S)$ in order to arrive at a three-dimensional smearing, then to compute the Poisson brackets of the regulated functions and finally we will remove the regulator and hope to arrive at a well-defined symplectic structure for the $A(p),E(S)$.

The simplest way to do this is to define a *tube* T^ϵ_p with central path p to be a smooth function of the form

$$T^{\epsilon t}_p:\ \mathbb{R}^{D-1}\times[0,1]\to\sigma;\ T^{\epsilon t}_p(s_1,.,s_{D-1},t'):=\delta^\epsilon(t'-t)\delta^\epsilon(s_1,..,s_{D-1})p_{s_1,..,s_{D-1}}(t') \quad \text{(I.3.1. 4)}$$

where $p_{s_1,..,s_{D-1}}$ is a smooth assignment of mutually non-intersecting paths diffeomorphic to $p:=p_{0,..,0}$ (a congruence) and δ^ϵ is a smooth regularization of the $\delta-$distribution in $\mathbb{R}^{D-1}$ and $\mathbb{R}$ respectively. We then define (recall formula (III.2.14) for the holonomy)

$$h^\epsilon_p(A):=\mathcal{P}e^{\int_{\mathbb{R}^{D-1}}d^{D-1}s\, f^\epsilon(s_1,..,s_{D-1})\int_0^1 dt\int_{p_{s_1,..,s_{D-1}}}\delta^\epsilon_t A} \quad \text{(I.3.1. 5)}$$

where path ordering is with respect the t parameter. We obviously have $\lim_{\epsilon\to 0}h_{T^\epsilon_p}=h_p$ pointwise in $\mathcal{A}$ for any choice of δ^ϵ. Likewise we define a *disk* D^ϵ_S with central surface S to be a smooth function of the form

$$D^\epsilon_S:\ \mathbb{R}\times U\to\sigma;\ D^\epsilon_p(s;u_1,.,u_{D-1}):=\delta^\epsilon(s)S_s(u_1,..,u_D) \quad \text{(I.3.1. 6)}$$

where S_s is a smooth assignment of mutually non-intersecting surfaces diffeomorphic to $S:=S_0$ (a congruence). Here U denotes the subset of $\mathbb{R}^{D-1}$ in the pre-image of S. We then define

$$E^\epsilon(S):=\int_{\mathbb{R}}ds\ \delta^\epsilon(s)E(S_s) \quad \text{(I.3.1. 7)}$$

We obviously have $\lim_{\epsilon\to 0}E(D^\epsilon_S)=E(S)$ pointwise in $\mathcal{E}$.

Next recall that the Poisson bracket algebra among the functions $F(A) = \int d^D x A_a^j F_j^a$, $E(f) = \int d^D x E_j^a f_a^j$ of section I.1 is isomorphic with a subalgebra of the Lie algebra $C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ of smooth functions and vector fields (derivatives on functions) on $\mathcal{A}$ respectively . This Lie algebra is defined by

$$[(\phi,\nu),(\phi',\nu')] := (\nu(\phi') - \nu'(\phi), [\nu,\nu']) \tag{I.3.1. 8}$$

where $\nu(\phi)$ denotes the action of the vector field ν on the function ϕ and $[\nu,\nu']$ denotes the Lie bracket of vector fields. The subalgebra of $C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ which is isomorphic to the Poisson subalgebra generated by the functions $F(A), E(f)$ is given by the elements $(F(A), E(f)) \mapsto (\phi_F, \beta\kappa\nu_f)$ with algebra

$$[(\phi_F,\nu_f),(\phi_{F'},\nu_{f'})] := (F'(f) - F(f'), 0) \tag{I.3.1. 9}$$

and if one would like to quantize the system based on the real-valued functions and vector fields ϕ_F, ν_f respectively, then one would ask to promote them to self-adjoint operators with commutator algebra isomorphic with (I.3.1. 9).

We are interested in quantizing the system based on another subalgebra of $C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ $(A(p), E(S) \mapsto (\phi_p, \beta\kappa\nu_S)$ which we now must derive using the above regularization. Let

$$\begin{aligned}
F_p^{\epsilon k t}(x)_j^a &:= \delta_j^k \int_{\mathbb{R}^{D-1}} d^{D-1}s\; \delta^\epsilon(s_1,..,s_{D-1}) \int_0^1 dt' \delta^\epsilon(t'-t) \dot{p}^a_{s_1,..,s_{D-1}}(t') \delta(x, p_{s_1,..,s_{D-1}}(t')) \\
f_S^{\epsilon k}(x)_a^j &:= \delta_j^k \int_{\mathbb{R}} ds\; \delta^\epsilon(s) \int_U d^{D-1}u\, \epsilon_{aa_1..a_{D-1}} \frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1} .. \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}} \times \\
&\quad \times \delta(x, S_s(u_1,..,u_{D-1}))
\end{aligned} \tag{I.3.1. 10}$$

then we trivially have

$$\begin{aligned}
h_p^\epsilon(A) &= \mathcal{P} e^{\int_0^1 F_p^{\epsilon j t}(A)\tau_j/2} \\
E_j^\epsilon(S) &= E(f_S^{\epsilon j})
\end{aligned} \tag{I.3.1. 11}$$

Notice that the smearing functions (I.3.1. 10) are not quite smooth due to the sharp cut-off at the boundary of the family of paths and surfaces respectively but this does not cause any trouble, the smeared functions are still differentiable because the functional derivatives (I.3.1. 10) define a bounded linear functional on $\mathcal{M}$ (see section I.1).

Formula (I.3.1. 11) enables us to map our regulated holonomy and surface variables into the Lie algebra $C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ via

$$h_p^\epsilon(A) \mapsto \phi_p^\epsilon := \mathcal{P} e^{\int_0^1 dt \phi_{F_p^{\epsilon j t}} \tau_j/2} \text{ and } E_j^\epsilon(S) \mapsto \nu_{Sj}^\epsilon := \nu_{f_S^{\epsilon j}} \tag{I.3.1. 12}$$

compute their algebra and then take the limit $\epsilon \to 0$ where we may use the known action of ν_f on ϕ_F.

Now the following issue arises:
By (I.3.1. 9) the vector fields $\nu_{f_S^{\epsilon j}}$ are Abelean at finite ϵ. On the other hand, we will compute a vector field ν_S by $\nu_S(\phi_p) := \lim_{\epsilon\to 0} \nu_{Sj}^\epsilon(\phi_p^\epsilon)$. But taking the limit $\epsilon \to 0$ and computing Lie brackets of vector fields might not commute, as we will see, it does not, the algebra of the ν_S will be non-Abelean, specifically, $[\nu_S, \nu_{S'}]$ is generically non-vanishing if $S \cap S' \neq \emptyset$. This is no cause of trouble because we will take the resulting limit Lie algebra as a starting point for quantization. It is here where it was important to have started with the Lie algebra of functions and vector fields, the commutator $[\nu_S, \nu_{S'}]$ is no longer of the form $\nu_{S''}$ and thus does not come from some $E(S'')$, hence, if we would

have based quantization on a Poisson algebra of functions we would get into trouble as it would not be a closed Poisson algebra of functions any longer. However, the Lie bracket algebra of vector fields is always closed and hence our elementary classical algebra that quantization will be based on will be the smallest closed subalgebra of $C^\infty(\mathcal{A})\times V^\infty(\mathcal{A})$ generated by the ϕ_p, ν_S. Of course, only vector fields of the form ν_S will have a classical interpretation as some $E(S)$ which is good enough in order to take the classical limit.

Let us then actually compute ϕ_p, ν_S:
The calculation is quite lengthy and involves expanding out carefully the path ordered exponential in (I.3.1. 12) and using the known action $\nu_f(\phi_F)=F(f)=\int d^DxF^a_j(x)f^j_a(x)$. We find

$$\begin{aligned}\nu^{\epsilon'}_{Sj}(\phi^\epsilon_p) &= \sum_{n=1}^\infty\int_0^1 dt_n\int_0^{t_n}dt_{n-1}..\int_0^{t_2}dt_1\sum_{k=1}^n\times\\ &\times (\phi_{F_p^{\epsilon j_1 t_1}}\tau_{j_1}/2)..(\phi_{F_p^{\epsilon j_{k-1} t_{k-1}}}\tau_{j_{k-1}}/2)[\nu_{f_S^{\epsilon' j}}(\phi_{F_p^{\epsilon j_k t_k}})\tau_{j_k}/2]\times\\ &\times (\phi_{F_p^{\epsilon j_{k+1} t_{k+1}}}\tau_{j_{k+1}}/2)..(\phi_{F_p^{\epsilon j_n t_n}}\tau_{j_n}/2)\end{aligned} \qquad \text{(I.3.1. 13)}$$

Using

$$\begin{aligned}\nu_{f_S^{\epsilon' j}}(\phi_{F_p^{\epsilon k t}}) &= \delta_{jk}\int_{\mathbb{R}^{D-1}}d^{D-1}s\,\delta^\epsilon(s_1,..,s_{D-1})\int_{\mathbb{R}}ds\,\delta^{\epsilon'}(s)\int_0^1dt'\delta^\epsilon(t'-t)\int_U d^{D-1}u\times\\ &\times \dot p^a_{s_1,..,s_{D-1}}(t')\epsilon_{aa_1..a_{D-1}}\frac{\partial S^{a_1}_s(u_1,..,u_{D-1})}{\partial u_1}..\frac{\partial S^{a_{D-1}}_s(u_1,..,u_{D-1})}{\partial u_{D-1}}\times\\ &\times \delta(S_s(u_1,..,u_{D-1}),p_{s_1,..,s_{D-1}}(t'))\end{aligned} \qquad \text{(I.3.1. 14)}$$

we can now take *first* the limit $\epsilon\to 0$ and *then* $\epsilon'\to 0$ (the reason for doing this will become transparent below). The result is

$$\begin{aligned}\nu^{\epsilon'}_{Sj}(\phi_p) &:= \sum_{n=1}^\infty\int_0^1 dt_n\int_0^{t_n}dt_{n-1}..\int_0^{t_2}dt_1\sum_{k=1}^n\times\\ &\times A(t_1)..A(t_{k-1})[\lim_{\epsilon\to 0}\nu_{f_S^{\epsilon' j}}(\phi_{F_p^{\epsilon j_k t_k}})\tau_{j_k}/2]A(t_{k+1})..A(t_n)\end{aligned} \qquad \text{(I.3.1. 15)}$$

where the limit in the square bracket is given by the distribution

$$\begin{aligned}&\delta_{jj_k}\int_{\mathbb{R}}ds\,\delta^{\epsilon'}(s)\int_U d^{D-1}u\,\dot p^a(t_k)\epsilon_{aa_1..a_{D-1}}\frac{\partial S^{a_1}_s(u_1,..,u_{D-1})}{\partial u_1}..\frac{\partial S^{a_{D-1}}_s(u_1,..,u_{D-1})}{\partial u_{D-1}}\times\\ &\times\delta(S_s(u_1,..,u_{D-1}),p(t_k))\end{aligned} \qquad \text{(I.3.1. 16)}$$

Luckily, there is an additional t_k integral involved in (I.3.1. 15) so that the end result will be non-distributional. Let $t\mapsto F(t)$ be any (integrable) function and consider the integral

$$\begin{aligned}&\int_{\mathbb{R}}ds\,\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_0^{t_{k+1}}dtF(t)\dot p^a(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S^{a_1}_s(u_1,..,u_{D-1})}{\partial u_1}..\frac{\partial S^{a_{D-1}}_s(u_1,..,u_{D-1})}{\partial u_{D-1}}\times\\ &\times\delta(S_s(u_1,..,u_{D-1}),p(t))\end{aligned} \qquad \text{(I.3.1. 17)}$$

Notice first of all that the derivative $\dot p$ is well-defined since p is piecewise analytic. Next, we can subdivide p into analytic segments e (edges) of the following four types:
up
e intersects S in one of its endpoints only, i.e. $q:=S\cap e=b(e)$ or $S\cap e=f(e)$ (but not both, subdivide an edge into two halves if necessary). Let $T_q(S)$ be the $D-1$ dimensional subspace of

the tangent space $T_q(\sigma)$ at q spanned by the vectors $\partial S/\partial u_k(u_1,..,u_{D-1})_{S(u)=q}$ tangential to S at q carrying the orientation induced from S, that is,

$$n_a(q) := \epsilon_{aa_1..a_{D-1}}(\frac{\partial S^{a_1}(u_1,..,u_{D-1})}{\partial u_1}..\frac{\partial S^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}})_{S(u)=q} \qquad (I.3.1.\ 18)$$

is the outward normal direction. Consider all derivatives $(d^n e/dt^n)(t)_{e(t)=q})$ and take the first one, $(d^{n_q}e/dt^{n_q})(t)_{e(t)=q}$ which, considered as a tangential vector, does not lie in $T_q(S)$. Then we require that

$$(-1)^{(n_q-1)\theta(q,e)}(d^{n_q}e^a/dt^{n_q})(t)_{e(t)=q}n_a(q) > 0 \qquad (I.3.1.\ 19)$$

where $\theta(q,e)=0$ if $q=b(e)$ and $\theta(q,e)=1$ if $q=f(e)$. If the orientation of e induced from p is such that it is outgoing from q, that is $q=b(e)$ then there is a neighbourhood U of q such that $U\cap r(e)$ lies "above" S. If e is ingoing, that is $q=f(e)$, then $q=b(e^{-1})$ so that e^{-1} is outgoing and we have $e^{(n)}(1)=(-1)^n(e^{-1})^{(n)}(0)$ where $e^{-1}(t)=e(1-t)$, so (I.3.1. 19) makes sure that $U\cap r(e^{-1})$ lies "below" S. Since $r(e)=r(e^{-1})$ also $U\cap r(e)$ lies below S in this case. One could summarize this by saying that the *up* case corresponds to edges whose orientation points "upwards" the normal direction of S.

down

The same as in the "up" case but now

$$(-1)^{(n_q-1)\theta(q,e)}(d^{n_q}e^a/dt^{n_q})(t)_{e(t)=q}n_a(q) < 0 \qquad (I.3.1.\ 20)$$

Now $U\cap r(e)$ lies "below" S if outgoing and "above" if ingoing, so the orientation of e is such that it points "downwards" the normal direction of S.

inside

The segment lies entirely inside S, that is $S\cap e=e$, so that for each $q\in e$ and any $n=1,2,..$ we have

$$(d^n e^a/dt^n)(t)_{e(t)=q}n_a(q) = 0 \qquad (I.3.1.\ 21)$$

outside

The segment e does not intersect S at all, that is, $e\cap S=\emptyset$.

First of all we notice that due to piecewise analyticity of p, p is a finite composition of entire analytic segments and for each of them the number of edges of the *up* and *down* type must be finite. Namely, otherwise we could draw an analytic, inextendable curve c within S (c is then analytic because it lies in the analytic surface S) through this infinite number of isolated intersection points which means that actually $S\cap e\subset c$ since e is analytic, that is, e has no isolated intersection points at all which is a contradiction. On the other hand, if e contains a segment of the *inside* type then e cannot have a segment of the *up* or *down* type because of analyticity, that is, $e=e_1\circ(e\cap U)\circ e_2$ where $e_1\cap U=e_2\cap U=\emptyset$ are of the *outside* type and $e\cap U$ is an (open, since U is open) segment of the *inside* type.

We conclude that p is a composition $p=e_1\circ e_2\circ..\circ e_N$ where each e_k is an analytic edge of a definite type. Let $e_l=p([t'_{l-1},t'_l])$ for some t'_l, $l=0,..,N$, $0=t'_0<t'_1<..<t'_N=1$ and define $0\le l(t)\le N-1$ to be the largest number such that $t'_{l(t)}\le t$. Then (I.3.1. 17) becomes

$$\sum_{l=1}^{l(t_{k+1})}\int_{\mathbb{R}} ds\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_{t'_{l-1}}^{t'_l} dtF(t)\dot{p}^a(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1}..$$

$$
\begin{aligned}
.. \quad & \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}}\delta(S_s(u_1,..,u_{D-1}),p(t)) \\
& + \int_{\mathbb{R}} ds\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_{t'_{l(t_{k+1})}}^{t_{k+1}} dtF(t)\dot{p}^a(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1}.. \\
.. \quad & \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}}\delta(S_s(u_1,..,u_{D-1}),p(t)) \\
= \quad & \sum_{l=1}^{l(t_{k-1})}\int_{\mathbb{R}} ds\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_0^1 dtF(\tilde{t}_l(t))\dot{e_l}^a(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1}.. \\
.. \quad & \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}}\delta(S_s(u_1,..,u_{D-1}),p(t)) \\
& + \int_{\mathbb{R}} ds\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_{t'_{l(t_{k+1})}}^{\delta_{l(t_{K+1})}(t_{K+1})} dtF(\tilde{t}_l(t))\dot{e}^a_{l(t_{k+1})+1}(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1}.. \\
.. \quad & \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}}\delta(S_s(u_1,..,u_{D-1}),p(t))
\end{aligned} \tag{I.3.1. 22}
$$

where in the second step we have used reparameterization invariance of (I.3.1. 17) and a reparameterization of e_l given by $[t'_{l-1},t'_l]=\tilde{t}_l([0,1])$ for $l\le l(t_{k-1})$ and $[t'_l,t_{k+1}]=\tilde{t}_l([0,\delta_{l(t_{k+1})}(t_{k+1})])$ for $l=l(t_{k+1})+1$ where $\delta_l(t)=0,1/2,1$ if $t=t'_{l-1},t'_{l-1}<t<t'_l,t=t'_l$.

Consider then for $t'=0,1/2,1$ the integral

$$
\begin{aligned}
& \int_{\mathbb{R}} ds\delta^{\epsilon'}(s)\int_U d^{D-1}u\int_0^{t'} dtF(t)\dot{e}^a(t)\epsilon_{aa_1..a_{D-1}}\frac{\partial S_s^{a_1}(u_1,..,u_{D-1})}{\partial u_1}.. \\
.. \quad & \frac{\partial S_s^{a_{D-1}}(u_1,..,u_{D-1})}{\partial u_{D-1}}\delta(S_s(u_1,..,u_{D-1}),e(t))
\end{aligned} \tag{I.3.1. 23}
$$

where $e\in\{e_1,..,e_N\}$ and $t'\in\{0,1/2,1\}$. We compute the value of (I.3.1. 23) for each possible type of e separately. Obviously, (I.3.1. 23) vanishes if $t'=0$ so we only consider $t'=1/2,1$.

Case *outside*:

This case is trivial, we let $\epsilon'\to 0$ in (I.3.1. 23) and see that the value of the integral becomes arbitrarily small and vanishes in the limit because the integrand has then support on an empty set.

Case *inside*

Since $s\mapsto S_s$ is a congruence it is clear that $\delta(S_s(u_1,..,u_{D-1}),e(t))$ has support at $s=0$ and the unique solution $u_1(t),..,u_{D-1}(t)$ (which are interior points of U since S is open) of the equation $S(u)=e(t)$. Thus (I.3.1. 23) becomes

$$
\delta^{\epsilon'}(0)\int_0^{t'} dtF(t)\frac{\dot{e}^a(t)\epsilon_{aa_1..a_{D-1}}[\frac{\partial S^{a_1}}{\partial u_1}..\frac{\partial S^{a_{D-1}}}{\partial u_{D-1}}]_{u(t)}}{|\det(\partial S_s(u)/\partial(s,u_1,..,u_{D-1}))_{s=0,u=u(t)}|} \tag{I.3.1. 24}
$$

which vanishes at finite ϵ' since the denominator is finite while the numerator vanishes by definition of an *inside* edge. Since (I.3.1. 24) vanishes at finite ϵ' its limit $\epsilon'\to 0$ vanishes as well. Expression (I.3.1. 24) is the precise reason for why we have not synchronized the limits $\epsilon\to 0,\epsilon'\to 0$ as otherwise we would have obtained an ill-defined result of the form $0\cdot\infty$.

Case *up*

Suppose first that $q=S\cap e=b(e)$. Expanding $e(t)$ around $t=0$ yields $e(t)=q+\frac{t^{n_q}}{n!}(d^{n_q}e/dt^{n_q})(0)+O(t^{n_q+1})$. Likewise let $u(q)$ be the unique solution of $S(u)=q$ and let us expand $S(u)=q+\sum_{k=1}^{D-1}(\partial S/\partial u_k)(u(q))[u_k-u(q)_k]+O((u-u(q))^2)$. Let us introduce new coordinates $v_k=u_k-$

$u(q)_k,\ k=1,..,D-1;\ v_D=t^n/(n!)$ and the matrix $(M_q)^a_k$ with $a,k=1,..,D$ and entries $M^a_k=(\partial S^a/\partial u_k)(u(q))$ for $k<D$ and $M^a_D=-(d^{n_q}e^a/dt^{n_q})(0)$. Letting $\epsilon'\to 0$, (I.3.1. 23) becomes

$$\begin{aligned}
& \int_U d^{D-1}u\int_0^{t'} dtF(t)[(d^ne^a/dt^{n_q})(0)t^{n_q-1}/((n_q-1)!)+O(t^{n_q})]\epsilon_{aa_1..a_{D-1}}\times\\
\times\ & [\frac{\partial S^{a_1}}{\partial u_1}(u(q))+O(v_1)]..[\frac{\partial S^{a_{D-1}}}{\partial u_{D-1}}(u(q))+O(v_{D-1})]\delta(M(q)\cdot v+O(v))\\
=\ & \int_{(U-u(q))\times[0,(t')^{n_q}/((n_q)!)]} d^DvF([(n_q!)v_D]^{1/n_q})(-\det(M_q)+O(v^{1/n_q}))\delta(M(q)\cdot v+O(v))\\
=\ & -\frac{\det(M_q)}{|\det(M_q)|}\frac{F(0)}{2}=\frac{F(0)}{2}
\end{aligned} \qquad \text{(I.3.1. 25)}$$

where in the last step we noticed that $-\det(M_q)=n_a(q)(e^{(n_q)})^a(0)>0$ by definition of the *up* type. The factor of $1/2$ is due to the fact that the support $v_D=0$ of the δ distribution is at the boundary of the v_D integral (while $v_k=0$ is in the interior of the v_k integral for $k<D$ since S is open), in other words, $\int_0^1 dt\delta(t)F(t)=F(0)/2$. One may ask whether we could not have chosen a different prescription for the value of the integral, say $\int_0^1 dt\delta(t)F(t)=sF(0)=F(0)-\int_{-1}^0 dt\delta(t)F(t)$ for some $1<s<1$. However, it is only for the value $s=1/2$ that the area operator, to be derived below, is invariant under switch of the orientation of the surface that it measures as it physically must be, see below. The reason is that now $\epsilon(S,e)=2s$ for the up type while $\epsilon(S,e)=-2(1-s)$ for the down type. Under switch of orientation of S this becomes $\epsilon(S,e)=2(1-s)$ for the up type while $\epsilon(S,e)=-2s$ for the up type.

Suppose now $q=S\cap e=f(e)$. Then, taking the limit $\epsilon'\to 0$ we see that (I.3.1. 23) vanishes if $t'=1/2$ while by a similar calculation it takes the value $F(1)/2$ if $t'=1$ (switch to $v_D=(1-t)^{n_q}/(n_q!)$ instead noticing that $e(t)=q+(-1)^{n_q}e^{(n_q)}(1)v_D+..,\ \dot{e}(t)=(-1)^{n_q-1}|\dot{v}_D|$).

We conclude that the value of the integral (I.3.1. 23) is given by $[F(0)\theta(t'-1/4)\delta_{S\cap e,b(e)}+F(1)\theta(t'-1)\delta_{S\cap e,f(e)}]/2$ where $\theta(t')=1$ for $t'\geq 0$ and $\theta(t')=0$ for $t'<0$ denotes the step function.
Case *down*
The calculation is completely analogous with the result that (I.3.1. 23) is given by $-[F(0)\theta(t'-1/4)\delta_{S\cap e,b(e)}+F(1)\theta(t'-1)\delta_{S\cap e,f(e)}]/2$.

We can summarize the analysis by defining $\epsilon(e,S)$ to be $+1,-1,0$ whenever e has type *up, down or in(out)side* respectively whence the value of (I.3.1. 23) is given by

$$\epsilon(e,S)[F(0)\theta(t'-1/4)\delta_{S\cap e,b(e)}+F(1)\theta(t'-1)\delta_{S\cap e,f(e)}]/2 \qquad \text{(I.3.1. 26)}$$

Inserting (I.3.1. 26) into (I.3.1. 15) we obtain

$$\begin{aligned}
& \nu_{Sj}(\phi_p):=\lim_{\epsilon'\to 0}\nu^{\epsilon'}_{Sj}(\phi_p)\\
=\ & \frac{1}{2}\sum_{n=1}^\infty\int_0^1 dt_n\int_0^{t_n}dt_{n-1}..\int_0^{t_2}dt_1\sum_{k=1}^n A(t_1)..A(t_{k-1})\times\\
\times\ & \{\sum_{l=1}^{l(t_{k+1})}\epsilon(e_l,S)[\delta(t_k-\tilde{t}_l(0))\delta_{S\cap e_l,b(e_l)}+\delta(t_k-\tilde{t}_l(1))\delta_{S\cap e_l,f(e_l)}]\\
& +\epsilon(e_{l(t_{k+1})+1},S)[\delta(t_k-\tilde{t}_{l(t_{k+1})+1}(0))\theta(\delta_{l(t_{k+1})}(t_{k+1})-1/4)\delta_{S\cap e_{l(t_{k+1})+1},b(e_{l(t_{k+1})+1})}\\
& +\delta(t_k-\tilde{t}_{l(t_{k+1})+1}(1))\theta(\delta_{l(t_{k+1})}(t_{k+1})-1)\delta_{S\cap e_{l(t_{k+1})+1},f(e_{l(t_{k+1})+1})}]\}\tau_j/2\times\\
& A(t_{k+1})..A(t_n)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{n=1}^{\infty}\int_0^1 dt_n\int_0^{t_n} dt_{n-1}..\int_0^{t_{k+2}} dt_{k+1}\sum_{k=1}^{n}\times \\
&\times \{\sum_{l=1}^{l(t_{k+1})}\epsilon(e_l,S)[\delta_{S\cap e_l,b(e_l)}\int_0^{\tilde{t}_l(0)} dt_{k-1}+\delta_{S\cap e_l,f(e_l)}\int_0^{\tilde{t}_l(1)} dt_{k-1}] \\
&+\epsilon(e_{l(t_{k+1})+1},S)[\theta(\delta_{l(t_{k+1})}(t_{k+1})-1/4)\delta_{S\cap e_{l(t_{k+1})+1},b(e_{l(t_{k+1})+1})}\int_0^{\tilde{t}_{l(t_{k+1})+1}(0)} dt_{k-1} \\
&+\theta(\delta_{l(t_{k+1})}(t_{k+1})-1)\delta_{S\cap e_{l(t_{k+1})+1},f(e_{l(t_{k+1})+1})}\int_0^{\tilde{t}_{l(t_{k+1})+1}(1)} dt_{k-1}]\}\times \\
&\times\int_0^{t_{k-1}} dt_{k-2}..\int_0^{t_2} dt_1 A(t_1)..A(t_{k-1})\tau_j/2A(t_{k+1})..A(t_n) \qquad \text{(I.3.1. 27)}
\end{aligned}
$$

Now, using the algebraic properties of the holonomy $h_{p\circ p'}=h_p h_{p'}$, $h_{p^{-1}}=h_p^{-1}$, working out the consequences in terms of path ordred exponentials, one can convince oneself after tedious algebra that

$$
\begin{aligned}
\nu_{Sj}(\phi_p) &= \sum_{l=1}^{N}\frac{\epsilon(e_l,S)}{2}h_{e_1}(A)..h_{e_{l-1}}(A)\times \\
&= [\delta_{e_l\cap S,b(e_l)}\frac{\tau_j}{2}h_{e_l}(A)+\delta_{e_l\cap S,f(e_l)}h_{e_l}(A)\frac{\tau_j}{2}]h_{e_{l+1}}(A)..h_{e_N}(A) \qquad \text{(I.3.1. 28)}
\end{aligned}
$$

This is our end result. Notice that the details of the regularization of the delta-distribtions did not play any role. It was seemingly important that we smeared via congruences of curves and surfaces as compared to more general smearings, however, any "reasonable" smearing admits a foliation via curves and surfaces respectively. Thus, the result (I.3.1. 28) is fairly general.

I.3.1.3 Invariant Vector Fields on G^n

The amazing feature of expression (I.3.1. 28) is that it is again a product of a finite number of holonomies, the harvest of having started from a manifestly background independent formulation. If we would have started from a function of E which is smeared in all D directions then this would be no longer true, (I.3.1. 28) would be replaced by a more complicated expression in which an additional integral over the extra dimension would appear.

The fact that (I.3.1. 28) is again a product of holonomies enables us to generalize the action of ν_{jS} to arbitrary cylindrical functions, restricted to smooth connections. Let $f\in \mathrm{Cyl}^1(\overline{\mathcal{A}})$, then we find a subgroupoid $l=l(\gamma)\in\mathcal{L}$ and $f_l\in C^1(X_l)$ such that $f=p_l^* f_l=[f_l]_\sim$ and a complex valued function F_l on $G^{|E(\gamma)|}$ such that $f(A)=f_l(p_l(A))=F_l(\rho_l(p_l(A)))$ with $\rho_l(A_l)=\{A_l(e)\}_{e\in E(\gamma)}=\{A(e)\}_{e\in E(\gamma)}$. We may choose γ in such a way that it is *adapted* to a given surface S, that is, each edge of γ has a definite type with respect to S. This will make the following computation simpler. Notice that every graph can be chosen to be adapted by subdividing edges appropriately. Let us now restrict f to $\mathcal{A}$ then

$$
[\nu_{Sj}(f)](A)=\frac{1}{2}\sum_{e\in E(\gamma)}\epsilon(e,S)[\delta_{e\cap S,b(e)}\frac{\tau_j}{2}A(e)+\delta_{e\cap S,f(e)}A(e)\frac{\tau_j}{2}]_{AB}\frac{\partial F_l}{\partial A(e)_{AB}}(\{A(e')\}_{e'\in E(\gamma)}) \quad \text{(I.3.1. 29)}
$$

Evidently, (I.3.1. 29) leaves $C^\infty(X_l)$ restricted to $\mathcal{A}$ invariant *which is why we can extend it to all of* $\overline{\mathcal{A}}$!

More precisely:
Define the so-called right – and left invariant vector fields on G by

$$\begin{aligned}(R_j f)(h) &:= (\frac{d}{dt})_{t=0} f(e^{t\tau_j}h) =: (\frac{d}{dt})_{t=0}[L^*_{e^{t\tau_j}}f](h) \\ (L_j f)(h) &:= (\frac{d}{dt})_{t=0} f(he^{t\tau_j}) =: (\frac{d}{dt})_{t=0}[R^*_{e^{t\tau_j}}f](h)\end{aligned} \qquad \text{(I.3.1. 30)}$$

where $R_h(h') = h'h,\ L_h(h') = hh'$ denotes the right and left action of G on itself. The right (left) invariance of R_j (L_j), that is, $(R_h)_* R_j = R_j$ ($(L_h)_* L_j = L_j$), follows immediately from the commutativity of left and right translations $L_h R_{h'} = R_{h'} L_h$. Notice, however, that the right invariant field generates left translations and vice versa. Then we can write (I.3.1. 29) in the compact form

$$Y^j_l(S)[f_l] = \frac{1}{4}\sum_{e\in E(\gamma)} \epsilon(e,S)[\delta_{e\cap S,b(e)} R^j_e + \delta_{e\cap S,f(e)} L^j_e] f_l \qquad \text{(I.3.1. 31)}$$

where R^j_e is R^j on the copy of G labelled by e and where from now on we just identify X_l with $G^{|E(\gamma)|}$ via ρ_l. Expression (I.3.1. 31) obviously does not require us to restrict $f = p^*_l f_l$ to $\mathcal{A}$ any more. Notice that while $Y^j_l(S)$, just as $E_j(S)$ does not have a simple transformation behaviour under gauge transformations, R^j_e, L^j_e in fact do

$$\begin{aligned}[(\lambda^e_g)^*([(\lambda^e_g)_* R^j_e](f_e))](h_e) &= [R^j_e((\lambda^e_g)^a st f_e)](h_e) = (\frac{d}{dt})_{t=0} f_e(g(b(e))e^{t\tau_j}h_e g(f(e))^{-1}) \\ &= (\frac{d}{dt})_{t=0} f_e(e^{t\mathrm{ad}_{g(b(e))}(\tau_j)} g(b(e))h_e g(f(e))^{-1}) \\ &= [(\lambda^e_g)^*(R^{\mathrm{ad}_{g(b(e))}(\tau_j)} f_e)](h_e)\end{aligned} \qquad \text{(I.3.1. 32)}$$

so that $(\lambda^e_g)_* R^j_e = [\mathrm{ad}_{g(b(e))}(\tau_j)]_k R^k_e$ where $\mathrm{ad}_{g(b(e))}(\tau_j) =: [\mathrm{ad}_{g(b(e))}(\tau_j)]_k \tau_k$. Similarly $(\lambda^e_g)_* L^j_e = [\mathrm{ad}_{g(f(e))}(\tau_j)]_k L^k_e$. This shows once more that R^j_e (L^j_e) is right (left) invariant.

We thus have found a family of vector fields $Y^j_l(S)$ whenever l is adapted to S. If $l = l(\gamma)$ is not adapted then we can produce an adapted one $l_S = l(\gamma')$ e.g. by choosing $r(\gamma) = r(\gamma')$ and by subdividing edges of γ into those with definite type with respect to S and where the edges of γ' carry the orientation induced by the edges of γ. Since $p^*_{l_S l} f_l \sim f_l$ we then simply *define*

$$p^*_{l_S l}(Y^j_l(S)(f_l)) := Y^j_{l_S}(p^*_{l_S l} f_l) \qquad \text{(I.3.1. 33)}$$

We must check that (I.3.1. 33) does not depend on the choice of an adapted subgroupoid. Hence, let l'_S be another adapted subgroupoid then we find $l_S, l'_S \prec l''_S$ which is still adapted (take for instance the union of the corresponding graphs and subdivide edges as necessary). Since (I.3.1. 33) is supposed to be a cylindrical function and $p_{L_s l} \circ p_{l''_S l_S} = p_{L_s l} \circ p_{l''_S l_S}$ we must show that

$$p^*_{l''_S l_S} Y^j_{l_S}(S)(p^*_{l_S l} f_l) = p^*_{l''_S l'_S} Y^j_{l'_S}(S)(p^*_{l'_S l} f_l) \qquad \text{(I.3.1. 34)}$$

As usual, if (I.3.1. 34) holds for one such adapted l''_S then it holds for all. To see that (I.3.1. 34) holds, it will be sufficient to show that for any adapted subgroupoids $l_S \prec l''_S$ we have

$$p^*_{l''_S l_S} Y^j_{l_S}(S)(f_{l_S}) = Y^j_{l''_S}(S)(p^*_{l''_S l_S} f_{l_S}) \qquad \text{(I.3.1. 35)}$$

from which then (I.3.1. 34) will follow due to $p_{l_s l} \circ p_{l''_S l_S} = p_{l'_s l} \circ p_{l''_S l'_S}$. We again need to check three cases:

a)
$e \in E(\gamma''_S)$ but $e \notin E(\gamma_S)$, then (I.3.1. 35) holds because $p^*_{l''_S l_S} f_{l_S}$ does not depend on $A(e)$ so that the additional terms proportional to R^j_e, L^j_e in (I.3.1. 31) drop out.
b)
$e \in E(\gamma''_S)$ but $e^{-1} \in E(\gamma_S)$. We observe that with $\tilde{f}(h) = f(h^{-1})$

$$(R^j \tilde{f})(h) = (\frac{d}{dt})_{t=0} f((e^{t\tau_j} h)^{-1}) = (\frac{d}{dt})_{t=0} f(h^{-1} e^{-t\tau_j}) = -(L^j f)(h^{-1}) \qquad \text{(I.3.1. 36)}$$

But then

$$\begin{aligned}
& \epsilon(e,S)[\delta_{e\cap S,b(e)} R^j_e + \delta_{e\cap S,f(e)} L^j_e]\tilde{f}_e \\
= & -\epsilon(e,S)[\delta_{e\cap S,b(e)} L^j_{e^{-1}} + \delta_{e\cap S,f(e)} R^j_{e^{-1}}] f_{e^{-1}} \\
= & \ \epsilon(e^{-1},S)[\delta_{e\cap S,b(e)} L^j_{e^{-1}} + \delta_{e\cap S,f(e)} R^j_{e^{-1}}] f_{e^{-1}} \\
= & \ \epsilon(e^{-1},S)[\delta_{e^{-1}\cap S,f(e^{-1})} L^j_{e^{-1}} + \delta_{e^{-1}\cap S,b(e^{-1})} R^j_{e^{-1}}] f_{e^{-1}} \qquad \text{(I.3.1. 37)}
\end{aligned}$$

where $\tilde{f}_e = p^*_{ee^{-1}} f_e$ in obvious notation. Hence (I.3.1. 35) is satisfied.
c)
$e_1, e_2 \in E(\gamma''_S)$ but $e = e_1 \circ e_2 \in E(\gamma_S)$. If $e \cap S = b(e)$ then $e_1 \cap S = b(e_1)$ and $\epsilon(e,S) = \epsilon(e_1,S)$ while $e_2 \cap S = \emptyset$ and $\epsilon(e_2,S) = 0$ (recall that $\epsilon(e,S) \neq 0$ implies that e,S intersect in only one point). Similarly, if $e \cap S = f(e)$ then $e_2 \cap S = f(e_2)$ and $\epsilon(e,S) = \epsilon(e_2,S)$ while $e_1 \cap S = \emptyset$ and $\epsilon(e_1,S) = 0$. Let $f_1(h_1) := f_2(h_2) = f(h_1 h_2)$ then due to left and right invariance

$$(R^j f_1)(h_1) = (R^j f)(h_1 h_2) \text{ and } (L^j f_2)(h_2) = (L^j f)(h_1 h_2) \qquad \text{(I.3.1. 38)}$$

hence

$$\begin{aligned}
& \sum_{I=1,2} \epsilon(e_I,S)[\delta_{e_I\cap S,b(e_I)} R^j_{e_I} + \delta_{e_I\cap S,f(e_I)} L^j_{e_I}] p^*_{e_1,e_2),e_1\circ e_2} f_e \\
= & \left\{ \begin{array}{ll} \epsilon(e_1,S) R^j_{e_1} p^*_{e_1,e_2),e_1\circ e_2} f_e & \text{if } e\cap S = b(e) \\ \epsilon(e_2,S) L^j_{e_2} p^*_{e_1,e_2),e_1\circ e_2} f_e & \text{if } e\cap S = f(e) \end{array} \right. \\
= & \left\{ \begin{array}{ll} \epsilon(e,S) R^j_e f_e & \text{if } e\cap S = b(e) \\ \epsilon(e,S) L^j_e f_e & \text{if } e\cap S = f(e) \end{array} \right. \\
= & \ \epsilon(e,S)[\delta_{e\cap S,b(e)} R^j_e + \delta_{e\cap S,f(e)} L^j_e] f_e \qquad \text{(I.3.1. 39)}
\end{aligned}$$

as claimed.

Hence our family of vector fields $(Y^j_l(S))_{l\in\mathcal{L}}$ is now defined for all possible $l \in \mathcal{L}$, in the language of the previous section we have the cofinal set $l_0 := l(\emptyset) \prec \mathcal{L}$. Let us check that it is a consistent family, that is

$$p^*_{l'l}(Y^j_l(S)(f_l)) = Y^j_{l'}(p^*_{l'l} f_l) \qquad \text{(I.3.1. 40)}$$

for all $l \prec l'$ which are not necessarily adapted. Given $l \prec l'$ we find always an adapted subgroupoid $l, l' \prec l_S$. Now by the just established independence on the adapted graph we may equivalently show that

$$p^*_{l_S l'} p^*_{l'l}(Y^j_l(S)(f_l)) = p^*_{l_S l'} Y^j_{l'}(p^*_{l'l} f_l) \qquad \text{(I.3.1. 41)}$$

Now since $p^*_{l_S l'} p^*_{l'l} = p^*_{l_S l}$ the left hand side equals $p^*_{l_S l}(Y^j_l(S)(f_l) \equiv Y_{l_S}(p^*_{l_S l} f_l)$ by definition of Y_l on arbitrary, not necessarily adapted graphs and the right hand side equals because of the same reason

$Y^j_{l_S}(p^*_{l_s l'} p^*_{l'l} f_l) = Y_{l_S}(p^*_{l_S l} f_l)$.

We thus have established that the family of vector fields $(Y^j_l(S))_{l\in\mathcal{L}}$ is a consistent family and defines a vector field $Y^j(S)$ on $\overline{\mathcal{A}}$. Notice moreover that $Y^j_l(S)$ is real valued: From (I.3.1. 31) this will follow if R_j, L_j are real valued. Now we have embedded G into a unitary group which means that $\bar{h}^T = h^{-1}$, in particular $\bar{\tau}^T_j = -\tau_j$. Hence

$$\begin{aligned} \overline{R^j_h} &= \overline{(\tau_j h)_{AB}} \partial/\partial \bar{h}_{AB} = -(h^{-1}\tau_j)_{BA} \partial/\partial \bar{h}^{-1}_{BA} \\ &= -(h^{-1}\tau_j)_{AB} (\partial h_{CD}/\partial \bar{h}^{-1}_{AB}) \partial/\partial h_{CD} = (h^{-1}\tau_j)_{AB} h_{CA} h_{BD} \partial/\partial h_{CD} = R^j_h \end{aligned} \quad \text{(I.3.1. 42)}$$

where use was made of $\delta h^{-1} = h^{-1} h h^{-1}$ and the fact that the symbol $\partial/\partial h_{AB}$ acts as if all components of h_{AB} were independent by definition of $R_j(f) = (\tau_j h)_{AB} \partial f/\partial h_{AB}$.

Next we consider its family of divergences with respect to the uniform measure μ_0. Now the projection μ_{0l} is simply the Haar measure on $G^{|E(\gamma)|}$. Since the Haar measure is right and left invariant, i.e. $(L_h)_*\mu_H = \mu_H = (R_h)_*\mu_H$ we have $\mathrm{div}_{\mu_H} R_j = \mathrm{div}_{\mu_H} L_j = 0$ as the following calculation shows:

$$-\int_G \mu_h [\mathrm{div}_{\mu_H} R_j] f = +\int_G \mu_H R_j(f) = (\frac{d}{dt})_{t=0} \int_G \mu_H L^*_{e^{t\tau_j}} f = (\frac{d}{dt})_{t=0} \int_G (L_{e^{t\tau_j}})_* \mu_H f = 0 \quad \text{(I.3.1. 43)}$$

It follows that $\mathrm{div}_{\mu_{0l}} Y^j_l(S) = 0$ so that $Y^j(S)$ is automatically μ_0 compatible (and the divergence is real valued).

I.3.1.4 Essential Self-Adjointness of Electric Flux Momentum Operators

Since Y^j_S is a consistently defined smooth vector field on $\overline{\mathcal{A}}$ which is μ_0-compatible, all the results from section I.2.4 with respect to the definition of corresponding momentum operators apply and the remaining question is whether the family of symmetric operators $P^j_l(S) := iY^j_l(S)$ with dense domain $D(P^j_l(S)) = C^1(X_l)$ is an essentially self-adjoint family.

Looking at (I.3.1. 31), essential self-adjointness of $P^j_l(S)$ on $L_2(X_l, d\mu_{0l})$ will follow if we can show that iR_j, iL_j are essentially self-adjoint on $L_2(G, d\mu_H)$ with core $C^1(G)$. That they are symmetric operators we know already. Now we we invoke the Peter&Weyl theorem that tells us that

$$L_2(G, d\mu_H) = \overline{\oplus_{\pi\in\Pi} L_2(G, d\mu_H)_{|\pi}} \quad \text{(I.3.1. 44)}$$

where Π is a collection of representatives of irreducible representations of G, one for each equivalence class, and $L_2(G, d\mu_H)_{|\pi}$ is the closed subspace of $L_2(G, d\mu_H)$ spanned by the matrix element functions $h \mapsto \pi_{mn}(h)$. The observation is now that R_j, L_j leave each $L_2(G, d\mu_H)_{|\pi}$ separately invariant. For instance

$$(R_j \pi_{mn})(h) = (\frac{d\pi_{mm'}(e^{t\tau_j})}{dt})_{t=0} \pi_{m'n}(h) \quad \text{(I.3.1. 45)}$$

It follows that iR_j, iL_j are symmetric operators on the finite dimensional Hilbert space $L_2(G, d\mu_H)_{|\pi}$ of dimension $\dim(\pi)^2$ and therefore are self-adjoint. Since the matrix element functions are smooth, by the basic criterion of essential self-adjointness it follows that $(i(R_j)_{|\pi} \pm i\cdot 1_\pi) C^\infty(G)_{|\pi}$ is dense in $L_2(G, d\mu_H)_{|\pi}$, hence so is $(i(R_j)_{|\pi} \pm i\cdot 1_\pi) C^1(G)_{|\pi}$. Correspondingly,

$$(iR_j \pm i\cdot 1) C^\infty(G) = \oplus_{\pi\in\Pi} (i(R_j)_{|\pi} \pm i\cdot 1_\pi) C^\infty(G)_{|\pi} \quad \text{(I.3.1. 46)}$$

is dense in $L_2(G, d\mu_H)$ and thus iR_j is essentially self-adjoint. The proof for iL_j is the same.

I.3.1.5 Selection of the Uniform Measure by Adjointness Conditions

We are now in the position to establish that $\mathcal{H}^0$ is a physically relevant Hilbert space. In fact, we ask the more general question whether this Hilbert space is in some sense naturally selected by just imposing the canonical commutation relations and the adjointness conditions:

The classical system is a Lie subalgebra of $C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ generated by pairs of the form ϕ_p, ν_{jS} associated with $A(p)$ and $E_j(S)$ where ϕ_p is G valued and ν_{jS} is real valued. We are therefore asked to find a representation of the Lie subagebra generated by the ϕ_p, ν_{js} by operators $\hat{A}(p), \hat{E}_j(S)$ on a Hilbert space $\mathcal{H}^0$ such that

$$(\hat{A}(p)_{AB})^\dagger = (\hat{A}(p)^{-1})_{BA} \text{ and } \hat{E}_j(S)^\dagger = \hat{E}_j(S) \tag{I.3.1. 47}$$

(remember that G is w.l.g. a subgroup of some $U(N)$). More precisely, $\hat{A}(p)_{AB}$ is supposed to be a bounded operator (no domain questions therefore) taking values in G and $\hat{E}_j(S)$ should be self-adjoint. Moreover, we must represent the bracket relations, that is the classical Lie algebra relations

$$[((\phi_p)_{AB}, \beta\kappa\nu_{jS}), ((\phi_{p'})_{A'B'}, \beta\kappa\nu_{j'S'})] = \beta\kappa(\nu_{jS}((\phi_{p'})_{A'B'}) - \nu_{j'S'}((\phi_p)_{AB}), \beta\kappa[\nu_{jS}, \nu_{j'S'}]) \tag{I.3.1. 48}$$

must be promoted to canonical commutation relations

$$\begin{aligned} &[((\hat{\phi}_p)_{AB}, \beta\kappa\hat{\nu}_{jS}), ((\hat{\phi}_{p'})_{A'B'}, \beta\kappa\hat{\nu}_{j'S'})] \\ =\;& i\beta\ell_p^2((\nu_{jS}((\phi_{p'})_{A'B'}))^\wedge - (\nu_{j'S'}((\phi_p)_{AB})))^\wedge, \beta\kappa([\nu_{jS}, \nu_{j'S'}])^\wedge) \end{aligned} \tag{I.3.1. 49}$$

where the Planck area $\ell_p^2 = \hbar\kappa$ has naturally come into play. Here $\hat{A}(p) := \phi_p$, $\beta\kappa\hat{\nu}_{jS} =: \hat{E}_j(S)$. It is clear that (I.3.1. 49) is trivially satisfied if we represent an element $(\phi, \nu) \in C^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$ on a Hilbert space of the form $L_2(\overline{\mathcal{A}}, d\mu)$ with some distributional extension $\overline{\mathcal{A}}$ of $\mathcal{A}$ and some measure μ thereon by

$$(\hat{\phi}\psi)(A) := \phi(A)\psi(A) \text{ and } (\hat{\nu}\psi)(A) := i\hbar[\nu[\psi](A) + (\phi_\nu\psi)(A)] \tag{I.3.1. 50}$$

where ϕ_ν is a linear function of ν, provided that the triple ϕ, ν, ϕ_ν can be extended from the smooth space $\mathcal{A}$ to the distributional space $\overline{\mathcal{A}}$. This will be true for our choice of $\overline{\mathcal{A}}, \phi_p, \nu_{jS}$ provided $\phi_{\nu_{jS}}$ can be chosen appropriately. So the canonical commutation relations are formally (since we did not discuss domain questions yet) satisfied then.

Now we come to the adjointness relations. Since $(\hat{\phi}_p)_{AB}$ is just a $G-$valued multiplication operator, the adjointness relation for G is trivially satisfied. Now ν_{jS} is real valued and in order to get $\hat{\nu}_{jS}$ symmetric to start with one should choose $\phi_{\nu_{jS}} = \frac{1}{2}\mathrm{div}_\mu\nu_{jS}$. Let now any measure μ be given and denote its push-forward to X_l by μ_l. Since X_l is finite dimensional, provided that μ_l is *absolutely continuous* with respect to μ_{0l}, there exists a mon-negative function ρ_l on X_l such that $\mu_l = \rho_l\mu_{0l}$. Since ν_{jS} leaves $C^\infty(X_l)$ invariant and has a push-forward $(\nu_{jS})_l$ to $C^\infty(X_l)$ given by (I.3.1. 31) which is a linear combination of left and right invariant vector fields, it follows that $\mathrm{div}_{\mu_l}(\nu_{jS})_l = (\nu_{jS})_l[\ln(\rho_l)]$. We therefore see that *the uniform measure* $\mu \equiv \mu_0$ *is uniquely picked* once we require $\mathrm{div}_\mu\nu_{jS} = 0$ and that μ is a probability measure which is regular with respect to μ_0. In other words, if we had not constructed μ_0 before, guided by diffeomorphism – and gauge invariance, we would have found this measure anyway now if we use the very natural condition of divergence freeness. Besides, it is not known whether there exists any other choice for μ such that the ν_{jS} are $\mu-$compatible.

Together with the former results we therefore arrive at the following classification theorem.

Theorem I.3.1
i)
Suppose that we want to base quantization on the classical Lie algebra $Cyl^\infty(\mathcal{A}) \times V^\infty(\mathcal{A})$*. Then the Lie subalgebra generated by holonomy functions and the vector fields corresponding to electric fluxes is well-defined and can be extended to a Lie subalgebra of* $Cyl^\infty(\overline{\mathcal{A}}) \times V^\infty(\overline{\mathcal{A}})$*.*
ii)
The electric flux vector fields arise from a consistent family of vector fields which are real-valued and compatible with the uniform measure, the corresponding divergence (which in fact vanishes) being real-valued. The corresponding momentum operator is essentially self-adjoint with respect to $\mathcal{H}^0 = L_2(\overline{\mathcal{A}}, d\mu_0)$ *with core* $Cyl^1(\overline{\mathcal{A}})$*.*
iii)
The uniform measure μ_0 *is uniquely selected among all regular, Borel probability measures* μ *on* $\overline{\mathcal{A}}$ *regular with respect to it by imposing that 1) the adjointness – and canonical commutation relations are implemented on* $L_2(\overline{\mathcal{A}}, d\mu)$ *and that 2) the electric flux vector fields are divergence free.*

I.3.2 Implementation of the Gauss Constraint

We do not really need to implement the Gauss constraint since we can directly work with gauge invariant functions (that is, one solves the constraint classically and quantizes only the phase space reduced with respect to the Gauss constraint). However, we will nevertheless show how to get to gauge invariant functions starting from gauge variant ones by using the technique of refined algebraic quantization outlined in section III.7.

I.3.2.1 Derivation of the Gauss Constraint Operator

We proceed similarly as in the case of the electric flux operator and start from the classical expression

$$G(\Lambda) := -\int d^D x [D_a \Lambda^j] E^a_j \equiv -E(D\Lambda) \tag{I.3.2. 1}$$

where $D_a\Lambda^j = \partial_a \Lambda^j + f^j{}_{kl} A^k_a \Lambda^l$ is the covariant derivative of the smearing field Λ^j. Notice that (I.3.2. 1) is almost an electric field smeared in D dimensions except that the smearing field $D\Lambda$ depends on the configuration space. Nevertheless the vector field on $\mathcal{A}$ corresponding to it is given by $-\kappa\beta\nu_{D\Lambda}$. Next we apply it to Cyl($\mathcal{A}$) by first computing its action on the special functions ϕ_p and then use the chain rule. In order to compute its action on ϕ_p we must regulate it as in the previous section and then define $\nu_{D\Lambda}(\phi_p) := \lim_{\epsilon\to 0} \nu_{D\Lambda}(\phi^\epsilon)$. Finally we hope that the end result is again a cylindrical function which we then may extend to $\overline{\mathcal{A}}$ and thus derive a cylindrical family of hopefully consistent vector fields on $\overline{\mathcal{A}}$.

We will not write all the steps, the details are precisely as in the previous section just that the additional limit $\epsilon' \to 0$ is missing. For the same reason a split of p into edges of different type is not necessary because E is smeared in D directions. One finds

$$\nu_{D\Lambda}(\phi_p) = \beta\kappa \int_0^1 dt \dot{p}^a(t) (D_a\Lambda^j)(p(t)) h_{p([0,t])}(A) \frac{\tau_j}{2} h_{p([t,1])}(A) \tag{I.3.2. 2}$$

Let us use the notation $\Lambda = \Lambda^j \tau_j$ and $A(p(t)) = \dot{p}^a(t) A^j_a(p(t))\tau_j/2$. Using $[\tau_j, \tau_k] = 2 f_{jk}{}^l \tau_l$ we can then recast (I.3.2. 2) into the form

$$\nu_{D\Lambda}(\phi_p) = \frac{\beta\kappa}{2} \int_0^1 dt h_{p([0,t])}(A) \{\frac{d}{dt}\Lambda(p(t)) + [A(p(t)), \Lambda(p(t))]\} h_{p([t,1])}(A) \tag{I.3.2. 3}$$

Now we invoke the parallel transport equation for the holonomy

$$\frac{d}{dt}h_{p([0,t])}(A) = h_{p([0,t])}(A) \tag{I.3.2. 4}$$

and use $h_{p([t,1])}(A) = h_{p([0,t])}(A)^{-1}h_p(A)$, then it is easy to see that (I.3.2. 3) becomes

$$\nu_{D\Lambda}(\phi_p) = \frac{\beta\kappa}{2}\int_0^1 dt\frac{d}{dt}\{h_{p([0,t])}(A)\Lambda(p(t))h_{(p([t,1])}(A)\} = \frac{\beta\kappa}{2}[-\Lambda(b(p))h_p(A) + h_p(A)\Lambda(f(p))] \tag{I.3.2. 5}$$

where we have performed an integration by parts in the last step. So indeed we *are* lucky: (I.3.2. 5) is a cylindrical function again. Let us write $\nu_\Lambda := -\nu_{D\Lambda}$ then for any $f_l \in C^\infty(X_l)$ for any subgroupoid $l = l(\gamma)$ we have

$$\begin{aligned}
[\nu_\Lambda(f_l)](A) &= \frac{\beta\kappa}{2}\sum_{e\in E(\gamma)}[\Lambda(b(e))A(e) - A(e)\Lambda(f(e))]_{AB}(\partial f_l/\partial A(e)_{AB})(A) \\
&= \frac{\beta\kappa}{2}\sum_{e\in E(\gamma)}([\Lambda_j(b(e))R^j_e - \Lambda_j(f(e))L^j_e]f_l)(A)
\end{aligned} \tag{I.3.2. 6}$$

Finally we write this as a sum over vertices in the compact form

$$G_l(\Lambda)[f_l] := \nu_\Lambda(f_l) = \frac{\beta\kappa}{2}\sum_{v\in V(\gamma)}\Lambda_j(v)[\sum_{e\in E(\gamma);\, v=b(e)} R^j_e - \sum_{e\in E(\gamma);\, v=f(e)} L^j_e]f_l \tag{I.3.2. 7}$$

Hence we have successfully derived a family of vector fields $G_l(\Lambda) \in V^\infty(X_l)$ for any $l \in \mathcal{L}$. No adaption of the graph was necessary this time. Since Λ_j is real valued for compact G, it follows from our previous analysis that $G_l(\Lambda)$ is real valued. Using the steps a), b) and c) of section I.3.1.3 one quickly verifies that it is a consistent family and that it is trivially μ_0-compatible because it is divergence-free since it is a linear combination of left – and right invariant vector fields. For the same reason, the associated momentum operator

$$\hat{G}_l(\Lambda)[f_l] == \frac{i\beta\ell_p^2}{2}\sum_{v\in V(\gamma)}\Lambda_j(v)[\sum_{e\in E(\gamma);\, v=b(e)} R^j_e - \sum_{e\in E(\gamma);\, v=f(e)} L^j_e]f_l \tag{I.3.2. 8}$$

is essentially self-adjoint with dense domain $C^1(\overline{\mathcal{A}})$.

I.3.2.2 Complete Solution of the Gauss Constraint

Using the Lie algebra of the left – and right invariant vector fields on X_l given by

$$[R^j_e, R^k_{e'}] = -2\delta_{ee'}f^{jk}\ _lR^l,\ \ [L^j, L^k] = 2\delta_{ee'}f^{jk}\ _lL^l,\ \ [R^j, L^k] = 0 \tag{I.3.2. 9}$$

(e.g. $([R^j, R^k]f)(h) = (\frac{\partial^2}{\partial s\partial s'})_{s=s'=0}f([e^{s'\tau_k}, e^{s\tau_j}]h)$) we find

$$\begin{aligned}
[G_l(\Lambda), G_l(\Lambda')] &= (\frac{\beta\kappa}{2})^2\sum_{e\in E(\gamma)}\{\Lambda_j(b(e))\Lambda'_k(b(e))[R^j_e, R^k_e] + \Lambda_j(f(e))\Lambda_k(f(e))[L^j_e, L^k_e]\} \\
&= -\beta\kappa G([\Lambda, \Lambda'])
\end{aligned} \tag{I.3.2. 10}$$

where we have defined $\Lambda(x) := \Lambda_j(x)\tau_j/2$. We see that the Lie algebra of the $G_l(\Lambda)$ represents the Lie algebra Lie(G) for each $l \in \mathcal{L}$ seprately and also represents the classical Poisson brackets among the Gauss constraints, see section I.1. This is already a strong hint that the condition $\hat{G}(\Lambda)=0$ for all smooth Λ_j really means imposing gauge invariance.

Let us see that this is indeed the case. According to the programme of RAQ we must choose a dense subspace of $\mathcal{H}^0$ which we choose to be $\mathcal{D}:=\mathrm{Cyl}^\infty(\overline{\mathcal{A}})$. Let $f=[f_l]_\sim$ be a smooth cylindrical function, that is, $f_l\in C^\infty(X_l)$, then $\hat{G}(\Lambda)f = p_l^a st(\hat{G}_l(\Lambda)f_l)$. We are looking for an algebraic distribution $L\in\mathcal{D}^*$ such that

$$L(p_l^*\hat{G}_l(\Lambda)f_l)=0 \tag{I.3.2. 11}$$

for all $\Lambda_j,\ l\in\mathcal{L},\ f_l\in C^\infty(X_l)$. Since, given l the smooth function Λ is still arbitrary, we may restrict its support to one of the vertices of γ with $l=l(\gamma)$ and see that (I.3.2. 11) is completely equivalent with

$$L(p_l^*[\sum_{e\in E(\gamma);\, v=b(e)} R_e^j - \sum_{e\in E(\gamma);\, v=f(e)} L_e^j]f_l)=0 \tag{I.3.2. 12}$$

for any $v\in V(\gamma),\ l\in\mathcal{L},\ f_l\in C^\infty(X_l)$.

We now use the fact that any function in $\mathcal{D}=C^\infty(\overline{\mathcal{A}})$ is a finite linear combination of spin-network functions T_s. Therefore, an element $L\in\mathcal{D}^*$ is completely specified by the complex values $L(T_s)$ with no growth condition on these complex numbers (an algebraic distribution is well-defined if it is defined pointwise in $\mathcal{D}$). We conclude that any element $L\in\mathcal{D}^*$ can be written in the form

$$L=\sum_{s\in\mathcal{S}} L_s < T_s,.> \tag{I.3.2. 13}$$

where $<.,.>$ denotes the inner product on $L_2(\overline{\mathcal{A}},d\mu_0)$ and $\mathcal{S}$ denotes the set of all spin-network labels. Now, first of all (I.3.2. 12) is therefore completely equivalent with

$$L(p_{l(\gamma(s))}^*[\sum_{e\in E(\gamma(s));\, v=b(e)} R_e^j - \sum_{e\in E(\gamma(s));\, v=f(e)} L_e^j]T_s)=0 \tag{I.3.2. 14}$$

for any $v\in V(\gamma(s)),\ s\in\mathcal{S}$ where $\gamma(s)$ is the graph that underlies s. Since the opertor involved in (I.3.2. 14) leaves $\gamma(s),\vec{\pi}(s)$ invariant and spin-network functions are mutually orthogonal we find that

$$\sum_{s'\in\mathcal{S},\,\gamma(s')=\gamma(s);\vec{\pi}(s')=\vec{\pi}(s)} L_{s'} < T_{s'},[\sum_{e\in E(\gamma(s));\, v=b(e)} R_e^j - \sum_{e\in E(\gamma(s));\, v=f(e)} L_e^j]T_s >=0 \tag{I.3.2. 15}$$

for any $v\in V(\gamma(s)),\ s\in\mathcal{S}$. Effectively the sum over s' is now reduced over all $\vec{m},\vec{n}$ with $m_e,n_e=1,..,d_{\pi_e}$ for any $e\in E(\gamma(s))$ and is therefore finite. From this it follows already that the most general solution L is an arbitrary linear combination of solutions of the form $<\psi,.>$ where ψ is actually normalizable.

Consider now an infinitesimal gauge transformation $g_t(x)=e^{t\Lambda_j(x)\tau_j}$ for some function $\Lambda_j(x)$ with $t\to 0$. Since $\overline{\mathcal{G}}\cong G^\sigma$ we may arrange that $g=1$ at all vertices of $\gamma(s)$ except for v. Our spin network function is of the form

$$T_s=[\prod_{e\in E(\gamma(s));\, b(e)=v} f_e(h_e)]\,[\prod_{e\in E(\gamma(s));\, f(e)=v} f_e(h_e)]F_s \tag{I.3.2. 16}$$

where F_s is a cylindrical function that does not depend on the edges incident at v. Then under an infinitesimal gauge transformation the spin-network function changes as

$$(\frac{d}{dt})_{t=0}\lambda_{g_t}^* T_s=(\frac{d}{dt})_{t=0}[\prod_{e\in E(\gamma(s));\, b(e)=v} f_e(g_t(v)h_e)]\,[\prod_{e\in E(\gamma(s));\, f(e)=v} f_e(h_e g_t(v)^{-1})]\ F_s$$

$$
\begin{aligned}
&= (\frac{d}{dt})_{t=0}[\circ_{e\in E(\gamma(s));\, b(e)=v}(L^e_{g_t(v)})^*] \circ [\circ_{e\in E(\gamma(s));\, f(e)=v}(R^e_{g_t(v)^{-1}})^*]\; T_s \\
&= \Lambda_j(v)[\sum_{e\in E(\gamma(s));\, b(e)=v} R^j_e - \sum_{e\in E(\gamma(s));\, f(e)=v} L^j_e]\; T_s \\
&= G_{l(\gamma(s))}(\Lambda)[T_s] \qquad \text{(I.3.2. 17)}
\end{aligned}
$$

which proves that $G_l(\Lambda)$ is the infinitesimal generator of $\lambda^l_{e^{t\Lambda}}$. It is therefore clear that the general solution L is a linear combination of solutions of the form $<\psi, .>$ where $\psi \in \mathcal{H}^0$ is gauge invariant. Strictly speaking, ψ has to be invariant under infinitesimal gauge transformations only but since G is connected there is no difference with requiring it to be invariant under all gauge transformations (the exponential map between Lie algebra and group is surjective since there is only one component, that of the identity).

We could therefore also have equivalently required that

$$L(\lambda^*_g f) = L(f) \qquad \text{(I.3.2. 18)}$$

for all $g \in \overline{\mathcal{G}}$ and all $f \in \mathcal{D} := C^\infty(\overline{\mathcal{A}})$. In passing we recall that we have defined in the previous section a unitary representation of $\overline{\mathcal{G}}$ on $\mathcal{H}^0$ defined densely on $C(\overline{\mathcal{A}})$ by $\hat{U}(g)f := \lambda^*_g$. Let $t \mapsto g_t$ be a continuous one-parameter subgroup of $\overline{\mathcal{G}}$, meaning that $\lim_{t\to 0} g_t(x) = g_0(x) \equiv 1_G$ for any $x \in \sigma$, meaning that $t \mapsto g_{tx} := g_t(x)$ is a continuous one parameter subgroup of G for any $x \in \sigma$ (if g_t is continuous at $t = 0$ then also at every s since $\lim_{t\to s} g_t = \lim_{t\to 0} g_t g_s = g_s$ since group multiplication is continuous). We claim that the one parameter subgroup of unitary operators $\hat{U}(t) := \hat{U}(g_t)$ is strongly continuous, that is, $\lim_{t\to 0} ||\hat{U}(t)\psi - \psi|| = 0$ for any $\psi \in \mathcal{H}^0$. Since any $\hat{U}(t)$ is bounded and $C^\infty(\overline{\mathcal{A}})$ is dense in $\mathcal{H}^0$ it will be sufficient to show that strong continuity holds when restricted to $\mathcal{D}$. Also, strong continuity follows already from weak continuity (i.e. $<\psi, \hat{U}(t)\psi'> \to <\psi, \psi'>$ for any $\psi, \psi' \in \mathcal{H}^0$) since $||\hat{U}(t)\psi - \psi||^2 = 2(||\psi||^2 - \Re(<\psi, \hat{U}(t)\psi>)$. Since $\mathcal{D}$ is spanned by finite linear combinations of mutually orthonormal spin network functions (they are in fact smooth), it will then be sufficient to show that $<T_s, \hat{U}(t)T_{s'}> \to <T_s, T_{s'}> = \delta_{ss'}$. If $s = (\gamma, \vec{\pi}, \vec{m}, \vec{n})$, $s' = (\gamma', \vec{\pi}', \vec{m}', \vec{n}')$ then a short computation, using that λ_g leaves $\gamma(s), \vec{\pi}(s)$ invariant, shows that

$$<T_s, \hat{U}(t)T_{s'}> = \delta_{\gamma,\gamma'}\delta_{\vec{\pi},\vec{\pi}'} \prod_{e\in E(\gamma)} [\pi_e(g_t(b(e)))_{m'_e m_e}\; \pi_e(g_t(f(e))^{-1})_{n_e n'_e} \qquad \text{(I.3.2. 19)}$$

and since the matrix element functions are smooth, the claim follows. We conclude therefore from Stone's theorem that for $g_t(x) = \exp(t\Lambda(x))$ the operator $\hat{G}(\Lambda)$ is the self-adjoint generator of $\hat{U}(t)$.

Finally we display the corresponding rigging map. Since $\overline{\mathcal{G}}$ is a group, the obvious ansatz is

$$\eta(f) := <\int_{\overline{\mathcal{G}}} \mu_H(g)\lambda^*_g f, .> \qquad \text{(I.3.2. 20)}$$

which, since λ^*_g preserves $C(C_l)$, is actually a map $\mathcal{D} \to \mathcal{D}$. Since μ_0 is a probability measure we could therefore immediately take the inner product on $\mathcal{H}^0$ for the solutions $\eta(f)$. But let us see where the rigging map proposal takes us. By definition

$$
\begin{aligned}
<\eta(f), \eta(f')>_\eta &:= \eta(f')[f] = \int_{\overline{\mathcal{G}}} \mu_H(g) <\lambda^*_g f, f'> \\
&= \int_{\overline{\mathcal{G}}} \mu_H(g) \int_{\overline{\mathcal{G}}} \mu_H(g') <\lambda^*_g f, \lambda_{g'} f'> = <\eta(f)^\dagger, \eta(f')^\dagger> \qquad \text{(I.3.2. 21)}
\end{aligned}
$$

where in the second equality we have observed that $<\lambda^*_g f, f'>$ is invariant under gauge transformations of f' and $\eta(f)^\dagger := <., \int_{\overline{\mathcal{G}}} \mu_H(g)\lambda^*_g f>$. So, indeed the gauge invariant inner product is just the restricted gauge variant inner product. Finally, for any gauge invariant observable we trivially have $\hat{O}'\eta(f) = \eta(\hat{O}f)$.

I.3.3 Implementation of the Diffeomorphism Constraint

Again we could just start from the fact that we have a reasonable unitary representation of the diffeomorphism group already defined in section I.2 but we wish to make the connection to the classical diffeomorphism constraint more clear in order to show that the representation defined really comes from the classical constraint. We will work at the gauge variant level in this section for convenience, however, we could immediately work at the gauge invariant level and all formulae in this section go through with obvious modifications. The reason for this is that the Gauss constraint not only forms a subalgebra in the full constraint algebra but actually an ideal, that is, since the Diffeomorphism and Hamiltonian constraint are actually gauge invariant, the corresponding operators leave the space of gauge invariant cylindrical functions invariant. Hence one can solve the Gauss constraint independently before or after solving the other two constraints.

I.3.3.1 Derivation of the Diffeomorphism Constraint Operator

The representation $\hat{U}(\varphi)$ of Diff(σ) was densely defined on spin network functions as

$$\begin{aligned}&\hat{U}(\varphi)T_s := T_{\varphi\cdot s} \text{ where } \varphi\cdot s := (\varphi\cdot e := \varphi^{-1}(e),\\&(\varphi\cdot\vec{\pi}(s))_{\varphi^{-1}(e)} := \pi_e, (\varphi\cdot\vec{m}(s))_{\varphi^{-1}(e)} := m_e, (\varphi\cdot\vec{n}(s))_{\varphi^{-1}(e)} := n_e)_{e\in E(\gamma(s))}\end{aligned} \tag{I.3.3. 1}$$

Let u be an analytic vector field on σ and consider the one parameter subgroup $t\to\varphi^u_t$ of Diff$^\omega(\sigma)$ (analytic diffeomorphisms) determined by the integral curves of u, that is, solutions to the differential equation $\dot{c}_{u,x}(t)=u(c(t))$, $c_{u,x}(0)=x$ with $\phi^u_t(x):=c_{u,x}(t)$. The classical diffeomorphism constraint is given by

$$V_a = H_a - A^j_a G_j = 2(\partial_{[a}A^j_{b]})E^b_j - A^j_a\partial_b E^b_j \tag{I.3.3. 2}$$

Smearing it with u gives

$$V(u) = \int d^3x(\mathcal{L}_u A^j)_a(x)E^a_j(x) = E(\mathcal{L}_u A) \tag{I.3.3. 3}$$

where $\mathcal{L}$ denotes the Lie derivative. Since the constraint is again linear in momenta we can associate with it a vector field $\beta\kappa\nu_{\mathcal{L}_u A}$ on $\mathcal{A}$ which again depends on A as well. Proceeding similarly as with the Gauss constraint we find for its action on holonomies of smooth connections

$$\nu_{\mathcal{L}_u A}\phi_p = \int_0^1 ds h_{p([0,s])}(A)(\mathcal{L}_u A)(p(s))h_{p([s,1])}(A) \tag{I.3.3. 4}$$

We claim that (I.3.3. 4) equals

$$(\frac{d}{dt})_{t=0}h_p((\varphi^u_t)^*A) \tag{I.3.3. 5}$$

To see this, one uses the expansion $(\varphi^u_t)^*A = A+t(\mathcal{L}_u A)+O(t^2)$ and the fact that with $p=p_1\circ..\circ p_N$ we have $h_p = h_{p_1}..h_{p_N}$ with $p_k=p([t_{k-1},t_k])$, $0=t_0<t_1<..<t_N=1$, $t_k-t_{k-1}=1/N$. Denote $\delta h_{p_k} := h_{p_k}(A+\delta A)-h_{p_k}(A)$ Hence

$$\begin{aligned}h_p(A+\delta A)-h_p(A) = &\sum_{n=1}^N\sum_{1\le k_1<..<k_n\le N}(h_{p_1\circ..\circ p_{k_1-1}}(A)[\delta h_{p_{k_1}}])(h_{p_{k_1+1}\circ..\circ p_{k_2-1}}(A)[\delta h_{p_{k_2}}])..\\&..(h_{p_{k_{n-1}+1}\circ..\circ p_{k_n-1}}(A)[\delta h_{p_{k_n}}])(h_{p_{k_n+1}\circ..\circ p_N}(A))\end{aligned} \tag{I.3.3. 6}$$

which holds at each finite N. Now using the formula $h_{p_k}(A)=\mathcal{P}\exp(A(p_k))$ where $A(p_k)=\int_{p_k}A^j\tau_j/2$ we obtain

$$\delta h_{p_k} = \mathcal{P}\{e^{[A+\delta A](p_k)} - e^{A(p_k)}\} \tag{I.3.3. 7}$$

so that δh_{p_k} is at least linear in δA and therefore in t for $\delta A = (\varphi_t^u)^* A - A$. Thus, dividing (I.3.3. 6) by t and taking the limit $t \to 0$ we find

$$(\frac{d}{dt})_{t=0} h_p((\varphi_t^u)^* A) = \sum_{k=1}^{N} h_{p_1 \circ .. \circ p_{k-1}}(A)[(\frac{d}{dt})_{t=0} h_{p_k}((\varphi_t^u)^* A)] h_{p_{k+1} \circ .. \circ p_N} \tag{I.3.3. 8}$$

Finally we have $h_{p_k}(A + \delta A) - h_{p_k}(A) = \delta A(p_k) + O(1/N^2)$ so that in the limit $t \to 0$ indeed (I.3.3. 8) turns into (I.3.3. 4).

Unfortunately, (I.3.3. 4) is no longer a cylindrical function and therefore we cannot construct a consistent family of cylindrically defined vector fields on $\overline{\mathcal{A}}$, in other words, (I.3.3. 4) cannot be extended to $\overline{\mathcal{A}}$. Of course for each s the functions $h_{p([0,s])}(A) = A(p([0, s])$ can directly be extended to $\overline{\mathcal{A}}$, however, $\mathcal{L}_u A$ makes only sense for smooth A. Moreover, we recall from section I.2 that the measure μ_0 is supported on connections A such that for any $p \in \mathcal{P}$ the function $s \mapsto A([0, s])$ is nowhere continuous and therefore unlikely to be mesurable with respect to ds. Thus, we are not able to define an operator that correspeonds to the infinitesimal diffeomorphism constraint.

The way out is the observation that *the action of finite diffeomorphisms* can be extended to $\overline{\mathcal{A}}$. In fact, the identity $\nu_{\mathcal{L}_u} h_p(A) = (\frac{d}{dt})_{t=0} h_p((\varphi_t^u)^* A)$ suggests to consider the exponentiation of the vector field $\nu_{\mathcal{L}_u A}$ which then gives the action $h_p(A) \mapsto h_p((\varphi_t^u)^* A)$. Since classically we can always recover the infinitesimal action from the exponentiated one, we do not lose any information. Moreover, we may consider general finite diffeomorphisms φ which unlike the φ_t^u are not necessarily connected to the identity. Now, by the duality between $p - chains$ and $p - forms$ we have for smooth A

$$h_p(\varphi^* A) = \mathcal{P} e^{\int_p \varphi^* A} = \mathcal{P} e^{\int_{\varphi^{-1}(p)} A} = h_{\varphi^{-1}(p)}(A) \tag{I.3.3. 9}$$

which is the reason for taking the inverse diffeomorphism in (I.3.3. 1). In the form (I.3.3. 9), it is clear that the finite action of $\mathrm{Diff}^\omega(\sigma)$ on $\mathcal{A}$ can be extended to $\overline{\mathcal{A}}$ when considering it as a map between homomorphisms. Hence

$$\delta: \ \mathrm{Diff}^\omega(\sigma) \times \overline{\mathcal{A}} \to \overline{\mathcal{A}}; \ \ (\varphi, A) \mapsto \delta_\varphi(A) \text{ where } [\delta_\varphi(A)](p) := A(\varphi^{-1}(p)) \tag{I.3.3. 10}$$

This furnishes the derivation of the action (I.3.3. 10) already defined in section I.2 from the classical diffeomorphism constraint. Notice that by bconstruction *the diffeomorphism quantum constraint algebra is free of anomalies*

$$\hat{U}(\varphi)\hat{U}(\varphi')\hat{U}(\varphi^{-1})\hat{U}((\varphi')^{-1}) = \hat{U}(\varphi \circ \varphi' \circ \varphi^{-1} \circ (\varphi')^{-1}) \tag{I.3.3. 11}$$

I.3.3.2 General Solution of the Diffeomorphism Constraint

We have seen that we can define a unitary representation of $\mathrm{Diff}^\omega(\sigma)$ on $\mathcal{H}^0$ by (I.3.3. 1) and that it is impossible to construct an action of the Lie algebra of $\mathrm{Diff}^\omega(\sigma)$ on $\overline{\mathcal{A}}$. We will now see that this has a counterpart for the representation $\hat{U}(\varphi)$: If there would be a quantum operator $\hat{V}(u)$ which generates infinitesimal diffeomorphisms, then it would be the self-adjoint generator of the one parameter subgroup $t \mapsto \hat{U}(\varphi_t^u)$, that is, we would have $\hat{U}(\varphi_t^u) = e^{it\hat{V}(u)}$. However, that generator exists only if the one parameter group is stronly continuous. We will now show that it is not strongly continuous. To see this, take any non-zero vector field and find an open subset $U \subset \sigma$ in which it is non-vanishing. We find a non-trivial graph γ contained in U and an infinite decreasing sequence (t_n) with limit 0 such that the graphs $\varphi^{-1}(\gamma)$ are mutually different. Take any spin network state T_s with $\gamma(s) = \gamma$. Since spin-network states over different graphs are orthogonal we have

$||\hat{U}(\varphi_0^u)T_s - T_s||^2 = 2$, thus proving our claim. This small computation demonstrates once again how distributional $\overline{\mathcal{A}}$ in fact is: Once a path just differs infinitesimally from a second one, they are algebraically independent and a distributional homomorphism is able to assign to them completely independent values, there is no continuity at all. This behaviour is drastically different from that of Gaussian measures and is deeply rooted in the background independence of our formalism: The covariance of a Gaussian measure depends on a background metric which is able to tell us *how far apart two points are*. However, in a diffeomorphism invariant theory there is no distinguished background metric, in contrast, there are diffeomorphisms which, with respect to *any* background metric, can take the two points as far apart or as close together as we desire, the positions of the two points are not gauge invariant.

The absence of an infinitesimal generator of diffeomorphisms is not necessarily bad because we can still impose diffeomorphism invariance via finite diffeomorphisms, in fact finite diffeomorphisms are even better suited to constructing a rigging map as we will see. However, it should be kept in mind that the passage from the connected component $\mathrm{Diff}_0^\omega(\sigma)$ to all of $\mathrm{Diff}^\omega(\sigma)$ is a non-trivial step which is not forced on us by the formalism. Since the so-called mapping class group $\mathrm{Diff}^\omega(\sigma)/\mathrm{Diff}_0^\omega(\sigma)$ is huge and not very well understood (see e.g. [148]), to take all of $\mathrm{Diff}^\omega(\sigma)$ is at least the most practical option then. Furthermore, one should stress once more that while analytic diffeomorphisms are not too bad (every smooth paracompact manifold admits a real analytic differentiable structure which is unique up to smooth diffeomorphisms, see e.g [149]) they are at least rather unnatural because the classical action has smooth diffeomorphisms as its symmetry group and also because an analytic diffeomorphism is determined already by its restriction to an arbitrarily small open subset U of σ. In particular, an analytic diffeomorphism cannot be the identity in U and non-trivial elsewhere. Hopefully these fine details will no longer be important in the final picture of the theory in which diffeomorphisms of any differentiability category should have at most a semiclassical meaning anyway.

Let us then go ahead and solve the finite diffeomorphism constraint. That is, by the methods of RAQ we are looking for algebraic distributions $L \in \mathcal{D}^*$ with $\mathcal{D} = C^\infty(\overline{\mathcal{A}})$ such that

$$L(\hat{U}(\varphi)f) = L(f) \;\forall \varphi \in \mathrm{Diff}^\omega(\sigma),\; f \in \mathcal{D} \tag{I.3.3. 12}$$

Here we have explicitly written out the invariance condition in terms of analytic diffeomorphisms. Since the span of spin network functions is dense in $\mathcal{D}$, (I.3.3. 12) is equivalent with

$$L(\hat{U}(\varphi)T_s) = L(T_s) \;\forall \varphi \in \mathrm{Diff}^\omega(\sigma),\; s \in \mathcal{S} \tag{I.3.3. 13}$$

In order to solve (I.3.3. 12) recall from section I.3.2.2 that every element of $\mathcal{D}^*$ can be written in the form $L = \sum_s L_s < T_s, . >$ where L_s are some complex numbers. Then (I.3.3. 13) becomes a very simple condition on the coefficients L_s given by

$$L_{\varphi \cdot s} = L_s \;\forall \varphi \in \mathrm{Diff}^\omega(\sigma),\; s \in \mathcal{S} \tag{I.3.3. 14}$$

Equation (I.3.3. 14) suggests to introduce the orbit $[s]$ of s given by

$$[s] = \{\varphi \cdot s;\; \varphi \in \mathrm{Diff}^\omega(\sigma)\} \tag{I.3.3. 15}$$

and therefore (I.3.3. 14) means that $s \mapsto L_s$ is constant on every orbit. Obviously, $\mathcal{S}$ is the disjoint union of orbits which motivates to introduce the space of orbits $\mathcal{N}$ whose elements we denote by ν. Introducing the elementary distributions $L_\nu := \sum_{s\in\nu} < T_s, . >$ we may write the general solution of the diffeomorphism constraint as

$$L = \sum_{\nu\in\mathcal{N}} c_\nu L_\nu \tag{I.3.3. 16}$$

for some complex coefficients c_ν which depend only on the orbit but not on the representative. Notice that $L_\nu(T_s) = \chi_\nu(s)$ where χ denotes the characteristic function.

We still do not have a rigging map but the structure of the solution space suggests to define

$$\eta(T_s) := \eta_{[s]} L_{[s]} \tag{I.3.3. 17}$$

for some complex numbers η_ν for each $\nu \in \mathcal{N}$ and to extend (I.3.3. 17) by linearity to all of $\mathcal{D}$, that is, one writes a given $f \in \mathcal{D}$ in the form $f = \sum_s f_s T_s$ with complex numbers $f_s = 0$ except for finitely many s and then defines $\eta(f) = \sum_s f_s \eta(T_s)$. This way the map η is tied to the spin network basis. The crucial question is now whether the coefficients can be chosen in such a way that η satisfies all requirements to be a rigging map.

Notice that η is almost an integral over the diffeomorphism group: One could have considered instead of η the following transformation

$$T_s \mapsto \sum_{\varphi \in \mathrm{Diff}^\omega(\sigma)} < \hat{U}(\varphi) T_s, . > \tag{I.3.3. 18}$$

and the right hand side is certainly diffeomorphism invariant. The measure that is being used here is a *counting measure* which is trivially translation invariant.

Unfortunately (I.3.3. 18) does not even define an element of $\mathcal{D}^*$ because there are uncountably infinitely many analytic diffeomorphisms which leave $\gamma(s)$ invariant. To see this, notice that if $e \in E(\gamma)$ is an analytic curve (edge) which is left invariant by φ then φ must leave invariant also the maximal analytic extension $\tilde{e}$ of e in σ due to analyticity (By definition, an analytic edge is given by D functions $t \mapsto e^a(t)$ so $\varphi_{|e}: \ t \mapsto \varphi(e(t))$ is again analytic which is why there must be an analytic reparameterization t_φ such that $\varphi(e(t)) = e(t_\varphi(t))$ if φ leaves e invariant. But then the range of the maximal analytic extension of $e \circ \tilde{t}_\varphi$ coincides with $\tilde{e}$ and equals $\varphi(\tilde{e})$). There is an analytic function $F_{\tilde{e}}$ which vanishes precisely on $\tilde{e}$ because the condition $F_{\tilde{e}}(\tilde{e}(t)) = 0$ for all t is equivalent to the condition that an analytic function on σ should vanish only on the x_1-axis in local coordinates, which is easily satisfied by the analytic function $x_2^2 + .. + x_D^2$ for instance. Hence, the function $F_{\tilde{\gamma}} := \prod_{e \in E(\gamma)} F_{\tilde{e}}$ vanishes precisely on $\tilde{\gamma}$. Thus, if we choose a constant vector field u on σ (which is trivially analytic) then $u_{\tilde{\gamma}} := F_{\tilde{\gamma}}^2 e^{-F_{\tilde{\gamma}}^2} u$ vanishes precisely on $\tilde{\gamma}$ and is analytic. It follows that its integral curves define a diffeomorphism which is trivial precisely on $\tilde{\gamma}$. Since there are uncountably many $u, F_{\tilde{\gamma}}$ the claim follows. As a consequence, (I.3.3. 18) contains uncountably many times the same functional $< T_s, . >$ so that its value on T_s diverges.

In a sense then, η is a group averaging map in which these *trivial action* diffeomorphisms have been factored out. Now while one can find a subgroup $\mathrm{Diff}^\omega_{[s]}(\sigma)$ of $\mathrm{Diff}^\omega(\sigma)$ such that

$$\eta(T_s) = \eta_{[s]} \sum_{\varphi \in \mathrm{Diff}^\omega_{[s]}(\sigma)} < \hat{U}(\varphi) T_s, . > \tag{I.3.3. 19}$$

(just choose precisely one diffeomorphism that maps s to a given $s' \in [s]$), unfortunately these subgroups depend on $[s]$ so that one cannot view (I.3.3. 19) as a regularized rigging map.

Let us see whether we can choose the coefficients $\eta_{[s]}$ in such a way that the rigging inner product is well-defined. By definition

$$< \eta(T_s), \eta(T_{s'}) >_\eta := \eta(T_s')[T_s] = \eta_{[s']} \chi_{[s']}(s) \tag{I.3.3. 20}$$

Thus, positivity requires that $\eta_{[s]} > 0$. Imposing hermiticity then requires that

$$\eta_{[s']} \chi_{[s']}(s) = \overline{< \eta(T_{s'}), \eta(T_s) >} = \overline{\eta(T_s)[T_{s'}]} = \eta_{[s]} \chi_{[s]}(s') \tag{I.3.3. 21}$$

Now both the right and left hand side are non vanishing if and only if $[s]=[s']$ so that (I.3.3. 21) is correct with no extra condition on the $\eta_{[s]}$.

Finally we come to the issue of diffeomorphism invariant observables. We call an operator $\hat{O}$ a strong observable if $\hat{U}(\varphi)\hat{O}\hat{U}(\varphi)^{-1}=\hat{O}$. We call it a weak observable if $\hat{O}'$ leaves the solution space invariant, in other words

$$\begin{aligned} L(\hat{U}(\varphi)f) &= L(f)\ \forall\,\varphi\in\text{Diff}^\omega(\sigma) \\ &\Rightarrow\ [\hat{O}'L](\hat{U}(\varphi)f)=L(\hat{O}^\dagger\hat{U}(\varphi)f)=L(\hat{U}(\varphi)^{-1}\hat{O}^\dagger\hat{U}(\varphi)f)=\hat{O}'L(f) \end{aligned} \tag{I.3.3. 22}$$

We now show that restricting attention to strong observables would lead to superselection sectors. Namely, suppose that $\hat{O}$ is a densely defined, closed, strongly diffeomorphism invariant operator and consider any two spin-network functions $T_s, T_{s'}$ with $\tilde{\gamma}(s)\neq\tilde{\gamma}(s')$ where $\tilde{\gamma}$ denotes the maximal analytic extension of γ. Then by the above construction we an at least countably infinite number of analytic diffeomorphisms φ_n with $\varphi_n(\gamma(s))=\gamma(s)$ but such that the $\varphi_n(\gamma(s'))$ are mutually different. Hence for any n

$$< T_{s'},\hat{O}T_s >=< T_{s'},\hat{U}(\varphi_n)^{-1}\hat{O}\hat{U}(\varphi_n)T_s >=<\hat{U}(\varphi_n)T_{s'},\hat{O}T_s > \tag{I.3.3. 23}$$

Since the states $\hat{U}(\varphi_n)T_{s'}$ are mutually orthogonal and since

$$||\hat{O}T_s||^2=\sum_{s''\in\mathcal{S}}|< T_{s''},\hat{O}T_s >|^2\geq\sum_{n=1}^\infty|<\hat{U}(\varphi_n)T_{s'},\hat{O}T_s >|=|< T_{s'},\hat{O}T_s >|^2\sum_{n=1}^\infty 1 \tag{I.3.3. 24}$$

we conclude that $< T_{s'},\hat{O}T_s >=0$. In other words, strongly diffeomorphism invariant, closed and densely defined operators cannot have matrix elements between spin network states defined over graphs with different maximal analytic extensions so that the Hilbert space would split into mutually orthogonal superselection sectors. If σ is compact, the total spatial volume would be an operator of that kind, it actually preserves the graph on which it acts. More generally, operators which are built entirely from electric field operators will have this property. However, classically the theory contains many diffeomorphism invariant functions which are not built entirely from electric fields but depend on the curvature of the connection (for instance the Hamiltonian constraint) and hence, as operators, do not leave the graph on which they act invariant (see the next section). Thus, it is not enough to consider only strongly invariant operators which is why no superselection takes place [60].

Next we show that there exists a choice for the $\eta_{[s]}$ such that $\hat{O}'\eta(f)=\eta(\hat{O}f)$ at least for strongly invariant operators which then by the general theory of section III.7 implies that the reality conditions $(\hat{O}')^\star=(\hat{O}^\dagger)'$ are satisfied. To choose the η_ν appropriately we must discuss the so-called symmetry group $P_{[s]}$ of $[s]$, defined as follows: Let p be a permutation on the set $E(\gamma(s))$ and define $p\cdot s:=(p(e),\pi_{p(e)}:=\pi_e, m_{p(e)}:=m_e, n_{p(e)}:=n_e)_{e\in E(\gamma)}$ (in the gauge invariant case a similar action is defined). Then $P_{[s]}$ is the subgroup of the permutation group consisting of those permutations such that for each $p\in P_{[s]}$ there exists an analytic diffeomorphism φ_p such that $\varphi_p\cdot s=p\cdot s$. It is clear that this definition is independent of the choice of the representative $s'\in[s]$. For instance, if γ is the figure eight loop (with intersection) and e,e' are its two edges then $P_{[s]}$ has two generators for $s=(\gamma,\pi_e=\pi_{e'},m_e=m_{e'},n_e=n_{e'})$ while there would be none if e.g. $\pi_e\neq\pi_{e'}$. This demonstrates that the orbit generating groups $\text{Diff}^\omega_{[s]}(\sigma)$ can have different sizes for $[s],[s']$ even if $\gamma(s),\gamma(s')$ are diffeomorphic.

Now we have just seen that a strong observable has matrix elements at most between $T_s, T_{s'}$ where $\tilde{\gamma}(s)=\tilde{\gamma}(s')$. The point is then the following: Let $[\gamma]=\{\varphi(\gamma);\ \varphi\in\text{Diff}^\omega(\sigma)\}$ be the orbit of a graph and let us consider a smaller subgroup $\text{Diff}^\omega_{[\gamma(s)]}(\sigma)$ of $\text{Diff}^\omega(\sigma)$ contained in $\text{Diff}^\omega_{[s]}(\sigma)$ and consisting of diffeomorphisms which map $\gamma(s)$ into precisely one of its orbit elements. Now, due to analyticity we can in fact choose $\mathrm{Diff}^\omega_{[\gamma]}(\sigma)=\mathrm{Diff}^\omega_{[\tilde\gamma]}(\sigma)$: If $\varphi\neq\varphi'\in\mathrm{Diff}^\omega_{[\gamma]}(\sigma)$ then $\varphi(\gamma)\neq\varphi'(\gamma)$ so certainly $\varphi(\tilde\gamma)\neq\varphi'(\tilde\gamma)$. Conversely if If $\varphi\neq\varphi'\in\mathrm{Diff}^\omega_{[\tilde\gamma]}(\sigma)$ then $\varphi^{-1}\circ\varphi'(\tilde s)\neq\tilde s$ for at least a segment $\tilde s$ of some edge of $\tilde\gamma$. However, $\tilde s$ belongs to the analytic extension of some edge e of γ. Suppose that $\varphi^{-1}\circ\varphi'(e)=e$. This is a contradiction because we have seen above that then $\varphi^{-1}\circ\varphi'$ preserves the whole analytic extension of e.

We conclude that the orbit size of $[s]$ is $|P_{[\tilde\gamma(s)]}|/|P_{[s]}|$ times the orbit size of $[\tilde\gamma(s)]$ where $P_{[\gamma]}$ is defined similarly as $P_{[s]}$ just that now $\varphi_p(\gamma)=p(\gamma)$ is required. (Again, if φ_p is a symmetry of γ then it is a symmetry of $\tilde\gamma$ by analyticity). We can therefore write

$$\eta(T_s)=\frac{\eta_{[s]}}{|P_{[s]}|}\sum_{\varphi\in\mathrm{Diff}^\omega_{[\tilde\gamma(s)]}(\sigma),\ p\in P_{[\tilde\gamma(s)]}}<\hat U(\varphi)\hat U(\varphi_p)T_s,.> \tag{I.3.3. 25}$$

Let now $\hat O$ be a strong observable then

$$\begin{aligned}<\eta(f),\hat O'\eta(T_s)>_\eta &= [\hat O'\eta(T_s)](f)=[\eta(T_s)](\hat O^\dagger f)\\ &= \frac{\eta_{[s]}}{|P_{[s]}|}\sum_{\varphi\in\mathrm{Diff}^\omega_{[\tilde\gamma(s)]}(\sigma),\ p\in P_{[\tilde\gamma(s)]}}<\hat U(\varphi)\hat U(\varphi_p)T_s,\hat O^d aggerf>\\ &= \frac{\eta_{[s]}}{|P_{[s]}|}\sum_{\varphi\in\mathrm{Diff}^\omega_{[\tilde\gamma(s)]}(\sigma),\ p\in P_{[\tilde\gamma(s)]}}<\hat U(\varphi)\hat U(\varphi_p)\hat O T_s,f>\\ &= <f,\eta(\hat O T_s)>_\eta \end{aligned} \tag{I.3.3. 26}$$

where in the last step we have used that $\hat O T_s$ is a countable linear combination of spin-network states $T_{s'}$ with $\tilde\gamma(s)=\tilde\gamma(s')$.

Hence there are in fact no additional conditions on $\eta_{[s]}$ as far as strong observables are concerned. If even the rigging map ansatz is general enough with respect to the weak observables is a completely different issue and not known at the moment. However, whether or not there is a rigging map with respect to the diffeomorphism constraint is of marginal interest anyway for the following reason: Remember that the classical constraint algebra between the Hamiltonian constraint $H(N)$ and Diffeomorphism constraint $\vec H(\vec N)$ respectively has the structure

$$\begin{aligned}&\{\vec H(\vec N),\vec H(\vec N')\}\propto\vec H([\vec N,\vec N']),\\ &\{\vec H(\vec N),H(N)\}\propto H(\vec N[N]),\ \{H(N),H(N')\}\propto\vec H(q^{-1}(NdN'-N'dN)\end{aligned} \tag{I.3.3. 27}$$

Thus, the Poisson Lie algebra of diffeomorphism constraints is actually a subalgebra (the first identity) of the full constraint algebra but it is not an ideal (the second identity). It is therefore not possible to solve the full constraint algebra in two steps by first solving the diffeomorphism constraint and then solving the Hamiltonian constraint in a second step: As (I.3.3. 27) shows, the dual Hamiltonian constraint operator cannot leave the space of diffeomorphism invariant distributions invariant and it is therefore meaningless to try to construct an inner product that solves only the diffeomorphism constraint. Rather, one has to construct the space of solutions of all constraints first before one can tackle the issue of the physical inner product.

I.4 Kinematical Geometrical Operators

In this section we will describe the so-called kinematical geometrical operators of Canonical Quantum Relativity. These are gauge invariant operators which measure the length, area and volume respectively of coordinate curves, surfaces and volumes for $D=3$. The area and volume operators were first considered by Rovelli and Smolin in the loop representation [150]. In [151] Loll divovered that the volume operator vanishes on gauge invariant states with at most trivalent vertices and used area and volume operators in her lattice theoretic framework [152]. Ashtekar and Lewandowski [153] used the connection representation defined in previous sections and could derive the full spectrum of the area operator while their volume operator differs from that of Rovelli and Smolin on graphs with vertices of valence higher than three which can be seen as the result of using different diffeomorphism classes of regularizations. In [154] de Pietri and Rovelli computed the matrix elements of the RS volume operator in the loop representation and de Pietri created a computer code for the actual case by case evaluation of the eigenvalues. In [147] the connection representation was used in order to obtain the complete set of matrix elements of the AL volume operator. Area and volume operator could be quantized using only the known quantizations of the electric flux of section I.3.1 but the construction of the length operator [156] required a new quantization technique which was actually first employed for the Hamiltonian constraint, see section II.1. To the same category of operators also belong the ADM energy surface integral [157], angle operators [158] and similar other operators that test components of the three metric tensor [159].

In D dimensions we have analogous objects corresponding to $d-$dimensional submanifolds of σ with $1 \le d \le D$. To get an idea of the constructions involved it will be sufficient here to describe the simplest operator, the so-called area operator which we construct in D dimensions and which measures the area of an open $D-1$ dimensional submanifold of σ. A common feature of all these operators is that they are essentially self-adjoint, positive semi-definite unbounded operators with pure point (discrete) spectrum which has a length, area, volume ... gap respectively of the order of the Planck length, area, volume etc. (that is, zero is not an accumulation point of the spectrum).

We call these operators kinematical because they do not (weakly) commute with the Diffeomorphism or Hamiltonian constraint operator. One may therefore ask what their physical significance should be. As a partial answer we will sketch a proof that if the curves, surfaces and regions are not coordinate manifolds but are invariantly defined through matter, then they not only weakly commute with the Diffeomorphism constraint but also their spectrum remains unaffected. There is no such argument with respect to the Hamiltonian constraint however. We will follow the treatment in [153].

I.4.1 Derivation of the Area Operator

Let S be an oriented, embedded, open, compactly supported, analytical surface and let $X : U \to S$ be the associated embedding where U is an open submanifold of $\mathbb{R}^{D-1}$. The area functional $\mathrm{Ar}[S]$ of the $D-$metric tensor q_{ab} is the volume of $X^{-1}(S)$ in the induced $(D-1)-$metric

$$\mathrm{Ar}[S] := \int_U d^{D-1}u \sqrt{\det([X^* q](u))} \tag{I.4.1. 1}$$

which coincides with the Nambu-Goto action for the bosonic Euclidean $(D-1)-$brane propagating in a $D-$dimensional target spacetime (σ, q_{ab}). Using the covector densities

$$n_a(u) := \epsilon_{a a_1 .. a_{D-1}} \prod_{k=1}^{D-1} \frac{\partial X^{a_k}}{\partial u_k}(u) \tag{I.4.1. 2}$$

familiar from section I.3.1 it is easy to see that we can write (I.4.1. 1 in the form

$$\mathrm{Ar}[S] := \int_U d^{D-1}u\sqrt{n_a(u)n_b(u)E^a_j(X(u))E^b_j(X(u))} \tag{I.4.1. 3}$$

Let now $U = \bigcup_{n=1}^N U'_n$ be a partition of U by closed sets U'_n with open interior U_n and let $\mathcal{U}$ be the collection of these open sets. Then the Riemann integral (I.4.1. 3) is the limit as $N \to \infty$ of the Riemann sum

$$\mathrm{Ar}_{\mathcal{U}}[S] := \sum_{U\in\mathcal{U}} \sqrt{E_j(S_U)E_j(S_U)} \tag{I.4.1. 4}$$

where $S_U = X(U)$ and $E_j(S)$ is the electric flux function of section I.3.1. The strategy for quantizing (I.4.1. 4) will be to use the known quantization of $E_j(S)$, to plug it into (I.4.1. 4), to apply it to cylindrical functions and to hope that in the limit $N \to \infty$ we obtain a consistently defined family of positive semi-definite operators. Notice that the square root involved makes sense because its argument will be a sum of squares of (essentially) self-adjoint operators which has non-negative real spectrum and we may therefore define the square root by the spectral resolution of the operator.

Let then $l = l(\gamma)$ be any subgroupoid and $f_l \in C^2(X_l)$. Using the results of section I.3.1 we obtain for any surface S

$$\hat{E}_j(S)\hat{E}_j(S)p^*_l f_l = -p^*_{l_S}\frac{\ell_p^4\beta^2}{16}\{\sum_{e\in E(\gamma_S)} \epsilon(e,S)[\delta_{e\cap S=b(e)}R^j_e + \delta_{e\cap S=f(e)}L^j_e]\}^2 p^*_{l_S l} f_l \tag{I.4.1. 5}$$

where $l_S = l(\gamma(S))$ is any adapted subgroupoid $l \prec l_S$.

When we now plug (I.4.1. 5) into (I.4.1. 4) we can exploit the following fact: Since (I.4.1. 4) classically approaches (I.4.1. 3) for *any* uniform refinement of the partition $\overline{\mathcal{U}}$, for given l and adapted l_S we can refine in such a way that for all $e \in E(\gamma)$ with $\epsilon(e,S) \neq 0$ (e is of the up or down type with respect to S) we have always that $e \cap S$ is an interior point of some $U \in \mathcal{U}$. Notice that then $\epsilon(e,S) = \epsilon(e,S_U)$ and $e\cap S = e \cap S_U$. If on the other hand $\epsilon(e,S) = 0$ but $S \cap e \neq \emptyset$ (e is of the inside type with respect to S) then for those U with $U \cap e \neq \emptyset$ we also have $\epsilon(e,S_U) = 0$. Clearly, if $e \cap S = \emptyset$ then $e \cap U = \emptyset$ for all $U \in \mathcal{U}$ so again $\epsilon(e,S) = \epsilon(e,S_U)$. We conclude that under such refinements the subgroupoid l_S stays adapted for all S_U. Let us denote an adapted partition and their refinements by $\mathcal{U}_l$. Then

$$\begin{aligned}\widehat{\mathrm{Ar}}_{\mathcal{U}_l}[S]p^*_l f_l &= \frac{\ell_p^2\beta}{4}p^*_{l_S}\sum_{U\in\mathcal{U}} \times \\ &\times\sqrt{-\{\sum_{e\in E(\gamma_S)} \epsilon(e,S_U)[\delta_{e\cap S_U=b(e)}R^j_e + \delta_{e\cap S_U=f(e)}L^j_e]\}^2} p^*_{l_S l} f_l \end{aligned}\tag{I.4.1. 6}$$

Let us introduce the set of isolated intersection points between γ and S

$$P_l(S) := \{e\cap S;\ \epsilon(e,S)\neq 0,\ e \in E(\gamma_S)\} \tag{I.4.1. 7}$$

which is independent of the choice of γ_S of course. After sufficient refinement, every S_U will contain at most one point which is the common intersection point of edges of the up or down type respectively. Let then for each $x \in P_l(S)$ the surface that contains x be denoted by S_{U_x}. From our previous discussion we know that then $\epsilon(e,S) = \epsilon(e,S_{U_x})$ for any $e \in E(\gamma_S)$ with $x \in \partial e$. It follows that (I.4.1. 6) simplifies after sufficient refinement to

$$\begin{aligned}\widehat{\mathrm{Ar}}_{\mathcal{U}_l}[S]p^*_l f_l &= \frac{\ell_p^2\beta}{4}p^*_{l_S}\sum_{x\in P_l(S)} \times \\ &\times\sqrt{-\{\sum_{e\in E(\gamma_S), x\in\partial e} \epsilon(e,S)[\delta_{x=b(e)}R^j_e + \delta_{x=b(e)}L^j_e]\}^2} p^*_{l_S l} f_l \end{aligned}\tag{I.4.1. 8}$$

Now the right hand side no longer depends on the degree of the adapted refinement and hence the limit becomes trivial

$$\widehat{\mathrm{Ar}}_{|l}[S]p_l^* f_l = \frac{\ell_p\beta}{4} p_{l_S}^* \sum_{x\in P_l(S)} \times \\ \times \sqrt{-\{\sum_{e\in E(\gamma_S), x\in\partial e} \epsilon(e,S)[\delta_{x=b(e)} R_e^j + \delta_{x=f(e)} L_e^j]\}^2} p_{l_S l}^* f_l \tag{I.4.1. 9}$$

Thus, we have managed to derive a family of operators $\widehat{\mathrm{Ar}}_l[S]$ with dense domain $\mathrm{Cyl}^2(\overline{\mathcal{A}})$. The independence of (I.4.1. 9) of the adapted graph follows from that of the $\hat{E}_j(S)$. Here we have encountered for the first time a common theme throughout the formalism: A state (or graph) dependent regularization. One must make sure therefore that the resulting family of operators is consistent.

I.4.2 Properties of the Area Operator

The following properties go through with minor modifications also for the length and volume operators.

1) *Consistency*
We must show that for any $l \prec l'$ holds that a) $\hat{U}_{ll'} C^2(X_l) \subset C^2(X_{l'})$ and that $\hat{U}_{ll'} \widehat{\mathrm{Ar}}_l[S] = \widehat{\mathrm{Ar}}_{l'}[S]\hat{U}_{ll'}$ where $\hat{U}_{ll'} f_l = p_{l'l}^* f_l$. Since the $p_{ll'}^*$ are analytic, a) is trivially satisfied. To verify b) we notice that (I.4.1. 9) can be written as

$$\widehat{\mathrm{Ar}}_{|l}[S] = \hat{U}_{l_S} \widehat{\mathrm{Ar}}_{l_S}[S] \hat{U}_{ll_S} \tag{I.4.2. 1}$$

where $\widehat{\mathrm{Ar}}_{l_S}[S]$ is simply the midlle operator in (I.4.1. 9) between the two pull-backs for the case that l is already adapted. First we must check that (I.4.2. 1) is independent of the adapted subgroupoid $l \prec l_S$. Let $l \prec l'_S$ be another subgroupoid and take a third adapted subgroupoid with $l_S, l'_S \prec l''_S$. If we can show that for any adapted subgroupoids with $l_S \prec l''_S$ we have

$$\widehat{\mathrm{Ar}}_{l''_S}[S]\hat{U}_{l_S l''_S} = \hat{U}_{l_S l''_S} \widehat{\mathrm{Ar}}_{l_S}[S] \tag{I.4.2. 2}$$

then we will be done. To verify (I.4.2. 2) we must make a case by case analysis as in section I.3.1.3 for the electric flux operator. But since (I.4.1. 9) is essentially the sum of square roots of the sum of squares of electric flux operators, the analysis is completely analogous and will not be repeated here.

Finally, let $l \prec l'$. We find an adapted subgroupoid $l, l' \prec l_S$. Then

$$\widehat{\mathrm{Ar}}_{|l'}[S]\hat{U}_{ll'} == \hat{U}_{l_S}\widehat{\mathrm{Ar}}_{l_S}[S]\hat{U}_{l'l_S}\hat{U}_{ll'} = \hat{U}_{l_S}\widehat{\mathrm{Ar}}_{l_S}[S]\hat{U}_{ll_S} = \widehat{\mathrm{Ar}}_{|l}[S] \tag{I.4.2. 3}$$

which is equivalent with consistency.
That the operator exists at all is like a small miracle: Not only did we multiply two functional derivatives $\hat{E}^a_j(x)$ at the same point, even worse, we took the square of it. Yet it is a densely defined, positive semi-definite operator without that we encounter any need for renormalization after taking the regulator (here the fineness of the partition) away. The reason for the existence of the operator is the *pay – off for having constructed a manifestly background independent representation.* We will see more examples of this "miracle" in the sequel.

2) *Essential Self-Adjointness*

To see that the area operator is symmetric, let $f_l \in C^2(X_l), f_{l'} \in C^2(X_{l'})$. Then we find an adapted subgroupoid $l, l' \prec l_S$ whence

$$\begin{aligned} < p_l^* f_l, \widehat{\mathrm{Ar}}[S] p_{l'}^* f_{l'} > &= < p_{l_S l}^* f_l, \widehat{\mathrm{Ar}}_{l_S}[S] p_{l_S l'}^* f_{l'} >_{L_2(X_{l_S}), d\mu_{0 l_S})} \\ &= < \widehat{\mathrm{Ar}}_{l_S}[S] p_{l_S l}^* f_l, p_{l_S l'}^* f_{l'} >_{L_2(X_{l_S}), d\mu_{0 l_S})} \\ &= < \widehat{\mathrm{Ar}}[S] p_l^* f_l, p_{l'}^* f_{l'} > \end{aligned} \tag{I.4.2. 4}$$

where in the second step we used that $\widehat{\mathrm{Ar}}_{l_S}[S]$ is symmetric on $L_2(X_{l_S}), d\mu_{0 l_S})$ with $C^2(X_{l_S})$ as dense domain.

Thus, the area operator is certainly a symmetric, positive semi-definite operator. Therefore we know that it possesses at least one self-adjoint extension, the so-called Friedrich's extension. However, we can show that $\widehat{\mathrm{Ar}}[S]$ is even essentially self-adjoint. The proof is quite similar to proving essential self-adjointness for the electric flux operator: Let $\mathcal{H}^0_{\gamma,\vec{\pi}}$ be the finite dimensional Hilbert subspace of $\mathcal{H}^0$ given by the closed linear span of spin-network functions over γ where all edges are labelled with the same irreducible representations given by $\vec{\pi}$. Then the Hilbert space maybe written as

$$\mathcal{H}^0 = \overline{\oplus_{\gamma \in \Gamma_0^\omega, \vec{\pi}} \mathcal{H}^0_{\gamma,\vec{\pi}}} \tag{I.4.2. 5}$$

Given a surface S we can without loss of generality restrict the sum over graphs to adapted ones because for $r(\gamma) = r(\gamma_S)$ we have $\mathcal{H}^0_{\gamma,\vec{\pi}} \subset \mathcal{H}^0_{\gamma_S,\vec{\pi}'}$ for the choice $\pi'_{e'} = \pi_e$ with $E(\gamma_S) \ni e' \subset e \in E(\gamma)$. Since then $\widehat{\mathrm{Ar}}[S]$ preserves each $\mathcal{H}^0_{\gamma,\vec{\pi}}$ its restriction is a symmetric operator on a finite dimensional Hilbert space, therefore it is self-adjoint. It follows that $\widehat{\mathrm{Ar}}_{|\gamma,\vec{\pi}}[S] \pm i \cdot 1_{\gamma,\vec{\pi}}$ has dense range on $\mathcal{H}^0_{\gamma,\vec{\pi}} = C^\infty(X_{l(\gamma)})_{\vec{\pi}} \subset C^2(X_{l(\gamma)})_{\vec{\pi}}$. Therefore

$$\begin{aligned} & [\widehat{\mathrm{Ar}}[S] \pm i \cdot 1_{\mathcal{H}^0}] C^2(\overline{\mathcal{A}}) = \oplus_{\gamma,\vec{\pi}} [\widehat{\mathrm{Ar}}_{|\gamma,\vec{\pi}}[S] \pm i \cdot 1_{\gamma,\vec{\pi}}] C^2(X_{l(\gamma)})_{\vec{\pi}} \\ & \supset \oplus_{\gamma,\vec{\pi}} [\widehat{\mathrm{Ar}}_{|\gamma,\vec{\pi}}[S] \pm i \cdot 1_{\gamma,\vec{\pi}}] \mathcal{H}^0_{\gamma,\vec{\pi}} = \oplus_{\gamma,\vec{\pi}} \mathcal{H}^0_{\gamma,\vec{\pi}} \end{aligned} \tag{I.4.2. 6}$$

is dense in $\mathcal{H}^0$.

3) *Spectral Properties*

i) *Discreteness*

Since $\widehat{\mathrm{Ar}}[S]$ leaves the $\mathcal{H}^0_{\gamma,\vec{\pi}}$ invariant it is simply a self-adjoint matrix there with non-negative eigenvalues. Since

$$\mathcal{H}^0_\gamma = \overline{\oplus_{\vec{\pi}} \mathcal{H}^0_{\gamma,\vec{\pi}}}$$

and the set of $\vec{\pi}$ is countable it follows that $\mathcal{H}^0_\gamma$ has a countable basis of eigenvectors for $\widehat{\mathrm{Ar}}[S]$ so that the spectrum is pure point (discrete), i.e. it does not have a continuous part. Now, as we vary γ we get a non-separable Hilbert space, however, the spectrum of $\widehat{\mathrm{Ar}}[S]$ depends only a) on the number of intersection points with edges of the up and down type, b) on their respective number per such intersection point and c) on the irreducible representations they carry and not on any other intersection characteristics. These possibilities are countable whence the entire spectrum is pure point and each eigenvalue comes with an uncountably infinite multiplicity.

ii) *Complete Spectrum*

It is even possible to compute the complete spectrum directly and to prove the discreteness from an explicit formula. Such a closed formula is unfortunately not available for the volume and length operator while badly needed for purposes in particular connected with quantum dynamics as we will see in the next section.

From the explicit formula (I.4.1. 9) it is clear that we may compute the eigenvalues for each intersection point x of S with edges of γ_S of the up or down type separately. Since the operator is independent of the choice of adapted graph, we may assume that all edges $e \in E(\gamma_S)$ have outgoing orientation, that is, $x = b(e)$ for each edge incident at x. Then (I.4.1. 9) reduces to

$$\widehat{\mathrm{Ar}}_{l_S}[S] = \frac{\ell_p^2\beta}{4} \sum_{x\in P_l(S)} \sqrt{-\{\sum_{e\in E(\gamma_S), x\in\partial e} \epsilon(e,S) R_e^j\}^2} \qquad \text{(I.4.2. 7)}$$

Let $E_{x,\star} = (\gamma_S) = \{e \in E(\gamma_S);\ x = b(e);\ e = \star\text{type}\}$ where $\star$ =u,d,i for "up, down, inside" respectively and let $R^j_{x,\star} = \sum_{e\in E_{x,\star}(\gamma_S)} R^j_e$. Then we have

$$\begin{aligned}\{\sum_{e\in E(\gamma_S), x\in\partial e} \epsilon(e,S) R_e^j\}^2 &= [R^j_{x,u} - R^j_{x,d}]^2 \qquad \text{(I.4.2. 8)}\\ &= (R^j_{x,u})^2 + (R^j_{x,d})^2 - 2R^j_u R^j_d = 2(R^j_{x,u})^2 + 2(R^j_{x,d})^2 - (R^j_u + R^j_d)^2\end{aligned}$$

where we have used that $[R^j_{x,u}, R^k_{x,d}] = 0$ (independent degrees of freedom). We check that $[R^j_{x,\star}, R^k_{x,\star}] = -2f_{jk}\,{}^l R^j_{x,\star}$ so that also $[R^j_{x,u+d}, R^k_{x,u+d}] = -2f_{jk}\,{}^l R^j_{x,u+d}$ with $R^j_{x,u+d} = R^j_u + R^j_v$. From this follows that $[R^k_\star, (R^j_u)^2] = [R^k_\star, (R^j_d)^2] = 0$ so that $\Delta_u = (R^j_{x,u})^2/4$, $\Delta_d = (R^j_{x,d})^2/4$, $\Delta_{u+d} = (R^j_{x,u+d})^2/4$ are mutually commuting operators and each of $R^j_{x,u}, R^j_{x,d}, R^j_{x,u+d}$ satisfies the Lie algebra of right invariant vector fields. Thus their respective spectrum is given by the eigenvalues $-\lambda_\pi < 0$ of the Laplacian $4\Delta = (R^j)^2 = (L^j)^2$ on G in irreducible representations π for which all matrix element functions π_{mn} are simultaneous eigenfunctions with the same eigenvalue. It follows that

$$\mathrm{Spec}(\widehat{\mathrm{Ar}}[S]) = \{\frac{\ell_p^2\beta}{2} \sum_{n=1}^N \sqrt{2\lambda_{\pi_n^1} + 2\lambda_{\pi_n^1} - \lambda_{\pi_n^{12}}};\ N \in \mathbb{N},\ \pi_n^1, \pi_n^2, \pi_n^{12} \in \Pi;\ \pi_n^{12} \in \pi_n^1 \otimes \pi_n^2\} \qquad \text{(I.4.2. 9)}$$

where the last condition means that π_n^{12} is an irreducible representation that appears in the decomposition into irreducibles of the tensor product representation $\pi_n^1 \otimes \pi_n^2$. In case that we are looking only at gauge invariant states we actually have $R^j_{x,u+v} = -R^j_{x,i}$. The spectrum (I.4.2. 9) is manifestly discrete by inspection. It is bounded from below by zero and is unbounded from above and depends explicitly on the Immirzi parameter.

iii) *Area Gap*

Let us discuss the spectrum more closely for $G = SU(2)$. Then per intersection point we have eigenvalues of the form

$$\lambda = \frac{\ell_p^2\beta}{2} \sqrt{2j_1(j_1+1) + 2j_2(j_2+1) - j_{12}(j_{12}+1)} \qquad \text{(I.4.2. 10)}$$

where $|j_1 - j_2| \le j_{12} \le j_1 + j_2$ by recoupling theory. Recoupling theory [160], that is, coupling of N angular momenta also tells us how to build the corresponding eigenfunctions through an appropriate recoupling scheme. The lowest positive eigenvalue is given by the minimum of (I.4.2. 10). At given j_1, j_2 the minimum is given at $j_{12} = j_1 + j_2$ which gives

$$\frac{\ell_p^2\beta}{2}\sqrt{(j_1-j_2)^2+j_1+j_2} = \frac{\ell_p^2\beta}{2}\sqrt{(j_2-(j_1-1/2))^2+2j_1-1/4} \tag{I.4.2. 11}$$

Since (I.4.2. 11) vanishes at $j_1 = j_2 = 0$ at least one of them must be greater than zero, say j_1. Then (I.4.2. 11) is minimized at $j_2 = j_1 - 1/2 \geq 0$ and proportional to $\sqrt{2j_1 - 1/4}$ which takes its minimum at $j_1 = 1/2$. Thus we arrive at the *area gap*

$$\lambda_0 = \frac{\sqrt{3}\ell_p^2\beta}{4} \tag{I.4.2. 12}$$

iv) *Main Series*

It is sometimes claimed [161] that the regularization of the area operator is incorrect and that a different regularization gives eigenvalues proportional to $\sqrt{j(j+1)}$ rather than (I.4.2. 10). If that would be the case then this would be of some significance for black hole physics as we will see in section II.3.2. However, first of all regularizations in quantum field theory are never unique and may lead to different answers, the only important thing is that all of them give the same classical limit. Secondly, even if the regularization performed in [161] is more aesthetic to some authors it is incomplete: In [161] one looks only at the so-called main series which results if we choose $j_1 = j_2 = j,\ j_{12} = 0$ and then just gives

$$\ell_p^2\beta\sqrt{j(j+1)}$$

(plus a quantum correction $j(j+1) \mapsto j(+1/2)^2$ due to the different regularization which results in integral quantum numbers). However, the complete spectrum (I.4.2. 10) is much richer, the side series have physical significance for the black hole spectrum as we will see and lead to a *correspondence principle*, that is, at large quantum numbers the spectrum approaches a continuum. To see this notice that at large eigenvalue λ changes as

$$\frac{\delta\lambda}{\lambda} \approx \frac{2(2j_1+1)\delta j_1 + 2(2j_2+1)\delta j_2 - (2j_{12}+1)\delta j_{12}}{2[(j_1+1)j_1 + (j_2+1)j_2 - (j_{12}+1)j_{12}]} \tag{I.4.2. 13}$$

Suppose we choose $j_1 = j_2 = j \gg 1$. Then $0 \leq j_{12} \leq 2j$ and we may choose $j_{12} = 0, \delta j_{12} = 1/2, \delta j_1 = \delta j_2 = 0$ (notice that such a transition is ignored if we do not discuss the side series). Then (I.4.2. 13) can be written

$$\delta\lambda \approx -\frac{(\lambda_0)^2}{\lambda} \tag{I.4.2. 14}$$

which becomes arbitrarily small at large j. The subsequent eigenvalues have been calculated numerically in [162] displaying a rapid transition to the continuum.

v) *Sensitivity to Topology*

The eigenvalues (I.4.2. 10) do detect some topological properties of σ as well. For instance, in the gauge invariant sector the spectrum depends on whether $\partial S = \emptyset$ or not. Moreover, for $\partial S = \emptyset$ the spectrum depends on whether S divides σ into two disjoint ergions or not.

I.4.3 Diffeomorphism Invariant Volume Operator

We now sketch how to make the geometrical operators at least a weak observable with respect to spatial diffeomorphisms. This is easiest for the volume functional.

Let R be a coordinate region, i.e. a $D-$dimensional submanifold of σ then the volume functional is defined by

$$\mathrm{Vol}[R] := \int_R d^D x \sqrt{\det(q)} = \int_\sigma d^D x \chi_R \sqrt{\det(q)} \qquad \text{(I.4.3. 1)}$$

where χ_R denotes the characteristic function of the set R. Suppose now that we couple gravity to matter (which is possible, see section II.2) and that ρ is a positive definite scalar density of any weight of the matter (and gravitational) degrees of freedom. Here by positive definite we mean that $\rho(x) = 0$ if and only if the matter field vanishes at x. For instance, if we have an electromagnetic field scalar field ϕ we could use the electromagnetic field energy density

$$\rho = \frac{q_{ab}}{2\sqrt{\det(q)}}[E^a E^b + B^a B^b]$$

Consider now the intrinsically defined region

$$R_\rho := \{x \in \sigma;\ \rho(x) > 0\} \qquad \text{(I.4.3. 2)}$$

Then

$$\mathrm{Vol}[R_\rho] = \int_\sigma d^D x \tilde{\theta}(\rho) \sqrt{\det(q)} \qquad \text{(I.4.3. 3)}$$

where $\tilde{\theta}$ is the modified step function with $\tilde{\theta}(x) = 1$ if $x > 0$ and $\tilde{\theta}(x) = 0$ otherwise. We claim that (I.4.3. 3) is in fact diffeomorphism invariant. To see this, it is sufficient to show that $F_\rho(x) := \tilde{\theta}(\rho(x))$ is a scalar of density weight zero. Let ρ be of density weight n, then under a diffeomorphism

$$F_\rho(x) \mapsto \tilde{\theta}(|\det(\partial\varphi(x)/\partial x)|^n \rho(\varphi(x))) = \tilde{\theta}(\rho(\varphi(x))) = (\varphi^* F_\rho)(x)$$

since $\tilde{\theta}(cx) = \tilde{\theta}(x)$ for any $c > 0$.

The use of matter is not really essential, we could also have used a gravitational degree of freedom say $\rho = \sqrt{\det(q)} R^2$ where R is the curvature scalar. The point is now that for scalar densities of weight one we can actually define $\hat{\rho}$ as an operator valued distribution (see section II.1) *if and only if ρ has density weight one*. Let $\mathcal{U}$ be a partion of σ. If it is fine enough and $\rho(x) > 0$ then also $\rho[U] := \int_U d^D x \rho(x) > 0$ for $x \in U \in \mathcal{U}$, therefore (I.4.3. 3) is approximated by

$$\mathrm{Vol}_{\mathcal{U}}[R_\rho] = \sum_{U \in \mathcal{U}} \tilde{\theta}(\rho[U]) \mathrm{Vol}[U] \qquad \text{(I.4.3. 4)}$$

Now $\rho[U]$ can be turned into a densely defined positive definite operator and thus $\tilde{\theta}(\hat{\rho}[U])$ can be defined by the spectral theorem. Moreover, since $\tilde{\theta}(x)^2 = \tilde{\theta}(x)$ we can order (I.4.3. 4) symmetrically and define

$$\widehat{\mathrm{Vol}}_{\mathcal{U}}[\rho] = \sum_{U \in \mathcal{U}} \tilde{\theta}(\rho[\hat{U}]) \widehat{\mathrm{Vol}}[U] \tilde{\theta}(\rho[\hat{U}]) \qquad \text{(I.4.3. 5)}$$

where for an adapted subgroupoid $l_U = l(\gamma_U)$

$$\widehat{\mathrm{Vol}}_{l_U}[U] = \frac{\beta^{3/2} \ell_p^3}{4} \sum_{v \in V(\gamma)} \sqrt{\left|\frac{1}{3!} \sum_{e,e',e'' \in E(\gamma_U);\ v=b(e)=b(e')=b(e'')} \epsilon(e, e'e'') f_{jkl} R^j_e R^k_{e'} R^l_{e''}\right|} \qquad \text{(I.4.3. 6)}$$

is the volume operator for coordinate regions. The adaption consists in orienting each edge to be outgoing from each vertex (for a given graph, subdivide each edge into two halves if necessary to get an adapted graph), the sum is over unordered triples of edges and

$$\epsilon(e, e', e'') = \mathrm{sgn}(\det(\dot{e}(0), \dot{e}'(0), \dot{e}''(0)))$$

The action on unadapted subgroupoids is defined similarly as for the area operator.

One now has to refine the partition and show that the final operator $\widehat{\mathrm{Vol}}[\rho]$, if it exists, is consistently defined. Since the spectrum of $\tilde{\theta}(\rho[\hat{U}])$ is given by $\{0, 1\}$, the spectra of that final operator and the coordinate volume operator should coincide and in that sense the discreteness of the spectrum is carried over to the diffeomorphism invariant context. Of course there remain technical issues, for instance $\widehat{\mathrm{Vol}}[U], \tilde{\theta}(\rho[\hat{U}])$ do not commute and cannot be diagonalized simultaneously, the existence of the limit is unclear etc. The details will appear elsewhere [163].

What this sketch shows are three points:
1) Kinematical Operators have a chance to become full Dirac observables by defining their coordinate regions invariantly through matter (for invariance under the Hamiltonian evolution, this requires them to be smeared over time intervals as well). Actually, this is physically the way that one defines regions !
2) The discreteness of the spectrum then has a chance to be an invariant property of the physical observables.
3) If true, then something amazing has happened:
We started out with an analytic manifold σ and smooth area functions. Yet, their spectra are entirely discrete, hinting at a discrete Planck scale physics, quantum geometry is distributional rather than smooth. Hopefully, the analytic structure that we needed at the classical level everywhere can be lifted to a purely combinatorial structure in the final picture of the theory, as it happened for $2+1$ gravity, see the fourth reference in [103].

Part II

Current Research

II.1 Quantum Dynamics

We now come to the "Holy Grail" of Canonical Quantum General Relativity, the definition of the Hamiltonian constraint. We will see that although one can, surprisingly, densely define a closed constraint operator at all, there is much less control on the correctness of the proposed operator than for the Gauss – and Diffeomorphism constraint operators. Also actually solving the proposed operator is not only technicly much more difficult but also conceptually: For instance, while RAQ gives some guidelines for how to do that and although it actually works (with some limitations) if we restrict ourselves to the spatial Diffeomorphism constraint as we have seen above, the definition of the physical inner product for all constraints, even for the already mentioned proposal, is an open problem so far. The reason is that the concept of a rigging map is currently out of control if the constraints do not form a Lie algebra as is the case for quantum gravity. Summarizing, the implementation of the correct quantum dynamics is not yet completed and one of the most active research directions at the moment.

While the situation with the proposed operator is certainly not completely satisfactory at the moment, in order to appreciate nevertheless its existence one should keep in mind that the situation with canonical quantum general relativity had *come to a sort of crisis in 1996*:

There were rigorous as well as formal results derived.

On the rigorous side one had constructed a *representation of the classical Poisson algebra for a suitable elementary set of "string-surface" variables, that is, a kinematical Hilbert space realized as an L_2 space with respect to a diffeomorphism invariant measure on a suitable quantum configuration space. Unfortunately, these results were not immediately useful for quantum gravity because the gravitational connection for the way the theory was defined at that time was complex valued rather than real valued and the kinematical Hilbert space defined above depends crucially on the fact that the connection is real-valued. It was considered impossible to quantize the density one valued unrescaled Hamiltonian constraint H in real variables because it is not polynomial.
This was the first big problem: The reality structure of Ashtekar's new variables had not been addressed yet, not even at the kinematical level.

On the formal side there were proposals for the quantization of the constraints and even for their kernel, however, the way they were defined was lacking diffeomorphism covariance, they included singular parameters and although they were meant for the complex Ashtekar connection, since the complex theory was not equipped with any Hilbert space it was unclear in which topology certain limits were performed and what the singular nature of the quantum field operators (and their products) was. For Euclidean gravity $\tilde{H}$ becomes actually polynomial in real variables but then one could show with the existing kinematical framework that the constraint operator was ill-defined in the given representation.
This was the second big problem: There existed no rigorous quantization of the constraints, especially not of the Wheeler-DeWitt constraint, all proposals were singular.

It seemed that one had a rigorous kinematical framework at one's disposal which was unphysical if one insisted in using complex variables (which was considerded mandatory) and which even in the unphysical representation did not support the Hamiltonian constraint operator !

There was some hope in terms of the Wick rotation proposal which we are going to sketch below which should keep the constraint polynomial and solve the reality conditions at the same time, however, that construction could be called at best formal and, moreover, the polynomial constraint for the real variables would suffer from the same singularities as the one for the complex variables.

The operator that we will describe below in principle kills both problems in one stroke: The crucial point was to realize that it is impossible to quantize the density weight two Hamiltonian constraint $\tilde{H}$ without breaking background independence. Could one then quantize the original density one valued Hamiltonian constraint H ? Since H is non-polynomial even in complex variables, the desire to have a complex connection formulation turned out to be of marginal interest. We then could show that with a new regularization technique, $\hat{H}$ can be turned into a well-defined operator using real-valued variables and using the established rigorous kinematical framework which now had become physically relevant. As a side result, also the generator of the Wick transform can be defined in principle using the same technique which could be a starting point for introducing the aesthetically more satisfactory complex variables into the framework again.

These considerations should be sufficient to indicate that the proposed operator, which we will describe in this section, is merely a first rigorous ansatz for the final operator but it is at least a promising hint that the kinematical framework that was developed *can support the Hamiltonian constraint operator*. It is arguably the most precisely defined ansatz that exists so far and hopefully it is a good starting point for improvements, generalizations and more drastic modifications (if necessary).

We will follow the only and exhaustive treatment in [80, 78, 164, 165, 119].

Remark:
Recently, a second approach towards solving the Hamiltonian constraint has been proposed [166, 167] which is constructed on (almost) diffeomorphism invariant distributions which are based on Vasiliev invariants. What is exciting about this is that one can define something like an area derivative [99] in this space and therefore the arc attachment which we will deal with exhaustively in what follows becomes much less ambiguous. In this review we will not describe this rather recent formalism because at the moment it falls outside of a Hilbert space context. Hopefully we can return to this in a future edition of this review when the theory has evolved more.

II.1.1 The Wick Transform Proposal

The Bargmann-Segal Transform for quantum gravity discussed in [83] gives a rigorous construction of quantum kinematics on a space of complexified, distributional connections by means of key results obtained by Hall [79]. Since the transform depended on a background structure, it was clear that the associated scalar product did not implement the correct reality conditions. To fix this was the purpose of [80] where a general theory was developed of how to trivialize reality conditions while keeping the algebraic structure of a functional as simple as when complex variables are being used. The same idea proves very useful in order to obtain a very general class of coherent states as we will see in section II.3. Moreover, as a side result, we were able to improve the coherent state transform as defined by Hall in the following sense :

Notice that the prescription given by Hall turns out to establish indeed a unitary transformation but that it was "pulled out of the hat", that is, it was guessed by an analogy consideration with the transform on $\mathbb{R}^n$ and turned out to work. It would be much more satisfactory to have a derivation of the transform $\hat{U}_t$ and the measure ν_t on the complexified configuration space from first principles, that is, one should be able to compute them just from the knowledge of the two polarizations of the phase space. We will first describe the general scheme in formal terms and then apply it to quantum gravity.

II.1.1.1 The General Scheme

Consider an arbitrary phase space $\mathcal{M}$, finite or infinite, with local real canonical coordinates (p, q) where q is a configuration variable and p its conjugate momentum (we suppress all discrete and continuous indices in this subsection). Furthermore, we have a Hamiltonian (constraint) $H'(p, q)$ which unfortunately looks rather complicated in the variables p, q (the reason for the prime will become evident in a moment). Suppose that, however, we are able to perform a *canonical* transformation on $\mathcal{M}$ which leads to the complex canonical pair $(p_{\mathbb{C}}, q_{\mathbb{C}})$ such that the Hamiltonian becomes algebraically simple (e.g.) a polynomial $H_{\mathbb{C}}$ in terms of $p_{\mathbb{C}}, q_{\mathbb{C}}$. That is, we have a complex symplectomorphism $(p_{\mathbb{C}}, q_{\mathbb{C}}) := W^{-1}(p, q)$ such that $H_{\mathbb{C}} = H' \circ W$ is algebraically simple. Notice that we are not complexifying the phase space, we just happen to find it convenient to coordinatize it by complex valued coordinates. The reality conditions on $p_{\mathbb{C}}, q_{\mathbb{C}}$ are encoded in the map W.

We now wish to quantize the system. We choose two Hilbert spaces, the first one, $\mathcal{H}$, for which the q's become a maximal set of mutually commuting, diagonal operators and a second one, $\mathcal{H}_{\mathbb{C}}$, for which the $q_{\mathbb{C}}$'s become a maximal set of mutually commuting, diagonal operators. According to the canonical commutation relations we represent $\hat{p}, \hat{q}$ on $\psi \in \mathcal{H}$ by $(\hat{p}\psi)(x) = i\hbar\partial\psi(x)/\partial x$ and $(\hat{q}\psi)(x) = x\psi(x)$. Likewise, we represent $\hat{p}_{\mathbb{C}}, \hat{q}_{\mathbb{C}}$ on $\psi_{\mathbb{C}} \in \mathcal{H}_{\mathbb{C}}$ by $(\hat{p}_{\mathbb{C}}\psi_{\mathbb{C}})(z) = i\hbar\partial\psi_{\mathbb{C}}(z)/\partial z$ and $(\hat{q}_{\mathbb{C}}\psi)(x) = z\psi_{\mathbb{C}}(z)$. The fact that p, q are real-valued force us to set $\mathcal{H} := L_2(\mathcal{C}, d\mu_0)$ where $\mathcal{C}$ is the quantum configuration space and μ_0 is the uniform (translation invariant) measure on $\mathcal{C}$ in order that $\hat{p}$ be self-adjoint.

In order to see what the Hilbert space $\mathcal{H}_{\mathbb{C}}$ should be, we also represent the operators $\hat{p}_{\mathbb{C}}, \hat{q}_{\mathbb{C}}$ on $\mathcal{H}$ by choosing a particular ordering of the function W^{-1} and substituting p, q by $\hat{p}, \hat{q}$. In order to avoid confusion, we will write them as $(\hat{p}', \hat{q}') := W^{-1}(\hat{p}, \hat{q})$ where the prime means that the operators are defined on $\mathcal{H}$ but are also quantizations of the classical functions $p_{\mathbb{C}}, q_{\mathbb{C}}$. Now, the point is that the operators $\hat{p}', \hat{q}'$, possibly up to $\hbar$ corrections, automatically satisfy the correct adjointness relations on $\mathcal{H}$ declining from the reality conditions on $p_{\mathbb{C}}, q_{\mathbb{C}}$. This follows simply by expanding the function W^{-1} in terms of $\hat{p}, \hat{q}$, computing the adjoint and defining the result to be the quantization of $\bar{p}_{\mathbb{C}}, \bar{q}_{\mathbb{C}}$ on $\mathcal{H}$ which equals any valid quantization prescription up to $\hbar$ corrections. Thus, if we could find a unitary operator $\hat{U} : \mathcal{H} \to \mathcal{H}_{\mathbb{C}}$ such that

$$\hat{p}_{\mathbb{C}} = \hat{U}\hat{p}'\hat{U}^{-1} \text{ and } \hat{q}_{\mathbb{C}} = \hat{U}\hat{q}'\hat{U}^{-1} \tag{II.1.1. 1}$$

then we have automatically implemented the reality conditions on $\mathcal{H}_{\mathbb{C}}$ as well because by unitarity

$$(\hat{p}_{\mathbb{C}})^\dagger = \hat{U}(\hat{p}')^\dagger\hat{U}^{-1} \text{ and } (\hat{q}_{\mathbb{C}})^\dagger = \hat{U}(\hat{q}')^\dagger\hat{U}^{-1} \tag{II.1.1. 2}$$

where the $\dagger$ operations in (II.1.1. 2) on the left and right hand side respectively are to be understood in terms of $\mathcal{H}_{\mathbb{C}}$ and $\mathcal{H}$ respectively. In other words, the adjoint of the operator on $\mathcal{H}_{\mathbb{C}}$ is the image of the correct adjoint of the operator on $\mathcal{H}$.

To see what $\hat{U}$ must be, let $\hat{K} : \mathcal{H} \cap \mathrm{Ana}(\mathcal{C}) \to \mathcal{H}_{\mathbb{C}}$ be the operator of analytical extension of real analytical elements of $\mathcal{H}$ and likewise $\hat{K}^{-1}$ the operator that restricts the elements of $\mathcal{H}_{\mathbb{C}}$ (all of which are holomorphic) to real values. We then have the identities

$$\hat{p}_{\mathbb{C}} = \hat{K}\hat{p}\hat{K}^{-1} \text{ and } \hat{q}_{\mathbb{C}} = \hat{K}\hat{q}\hat{K}^{-1} \ . \tag{II.1.1. 3}$$

We now exploit that W^{-1} was supposed to be a canonical transformation (an automorphism of the phase space that preserves the symplectic structure but not the reality structure). Let C be its infinitesimal generator, called the *complexifier*, that is, for any function f on $\mathcal{M}$,

$$f(p_{\mathbb{C}}, q_{\mathbb{C}}) := f_{\mathbb{C}}(p,q) := ((W^{-1})^* f)(p,q) = \sum_{n=0}^{\infty} \frac{i^n}{n!}\{C, f\}_{(n)} \tag{II.1.1. 4}$$

where the multiple Poisson bracket is inductively defined by $\{C,f\}_{(0)} = f$ and $\{C,f\}_{(n+1)} = \{C, \{C,f\}_{(n)}\}$. Using the substitution rule that Poisson brackets become commutators times $1/(i\hbar)$ we can quantize (II.1.1. 4) by

$$\hat{f}' := f_{\mathbb{C}}(\hat{p},\hat{q}) := \sum_{n=0}^{\infty} \frac{1}{\hbar^n n!}[\hat{C}, \hat{f}]_{(n)} = (\hat{W}_t)^{-1}\hat{f}\hat{W}_t \tag{II.1.1. 5}$$

where we have defined the generalized "heat kernel" operator

$$\hat{W}_t := e^{-t\hat{C}} \tag{II.1.1. 6}$$

and $t = 1/\hbar$. That is, the generator C motivates a natural ordering of $W^{-1}(p,q)$.

Substituting (II.1.1. 6) into (II.1.1. 4) we find

$$\hat{p}_{\mathbb{C}} = \hat{U}_t\hat{p}'\hat{U}_t^{-1} \text{ and } \hat{q}_{\mathbb{C}} = \hat{U}_t\hat{q}'\hat{U}_t^{-1} \tag{II.1.1. 7}$$

where we have defined the generalized coherent state or *Wick rotation transform*

$$\hat{U}_t := \hat{K}\hat{W}_t \tag{II.1.1. 8}$$

with $t = 1/\hbar$. The reason for the names we chose will become obvious in the next subsection.

It follows that if $\hat{C}, \hat{W}_t$ exist on real analytic functions and if we can then extend $\hat{U}_t$ to a unitary operator from $\mathcal{H}$ to $\mathcal{H}_{\mathbb{C}} := L_2(\mathcal{C}_{\mathbb{C}}, d\nu_t) \cap \mathrm{Hol}(\mathcal{C}_{\mathbb{C}})$ where $\mathcal{C}_{\mathbb{C}}$ denotes the complexification of $\mathcal{C}$ then we have completed the programme.

Moreover, as a bonus we would have *simplified the spectral analysis* of the operator that corresponds to the quantization of H' :
First of all we define an unphysical Hamiltonian (constraint) operator $\hat{H}$ on $\mathcal{H}$ simply by choosing a suitable ordering of the function

$$H(p,q) := H_{\mathbb{C}}(p_{\mathbb{C}}, q_{\mathbb{C}})_{|p_{\mathbb{C}} \to p, q_{\mathbb{C}} \to q} = (K^{-1} \cdot H_{\mathbb{C}})(p,q) \tag{II.1.1. 9}$$

and substituting p,q by the operators $\hat{p},\hat{q}$. Thus we obtain an operator $\hat{H}_{\mathbb{C}}$ on $\mathcal{H}_{\mathbb{C}}$ by $\hat{H}_{\mathbb{C}} := \hat{K}\hat{H}\hat{K}^{-1}$. It follows that if we *define* the quantization of the physical Hamiltonian (constraint) H' on $\mathcal{H}$ by $\hat{H}' := \hat{W}_t^{-1}\hat{H}\hat{W}_t$ then in fact $\hat{H}_{\mathbb{C}} = \hat{U}_t\hat{H}'\hat{U}_t^{-1}$ and since $\hat{U}_t$ is unitary the spectra of $\hat{H}'$ on $\mathcal{H}$ and of $\hat{H}_{\mathbb{C}}$ on $\mathcal{H}_{\mathbb{C}}$ *coincide*. But since $\hat{H}_{\mathbb{C}}$ is an algebraically simple function of the elementary operators $\hat{p}_{\mathbb{C}}, \hat{q}_{\mathbb{C}}$ it follows that one has drastically simplified the spectral analysis of the complicated operator $\hat{H}'$! Finally, given a (generalized) eigenstate $\psi_{\mathbb{C}}$ of $\hat{H}_{\mathbb{C}}$, we obtain a (generalized) eigenstate $\psi := \hat{U}_t^{-1}\psi_{\mathbb{C}}$ of $\hat{H}'$ by the *inverse* of the coherent state transform.

The crucial question then is whether we can actually make $\hat{U}_t$ unitary. In [80] we derived the following formula for the unitarity implementing measure ν_t on $\mathcal{C}_{\mathbb{C}}$:

$$\begin{aligned} d\nu_t(z,\bar{z}) &:= \nu_t(z,\bar{z})d\mu_0^{\mathbb{C}}(z)\otimes d\bar{\mu}_0^{\mathbb{C}}(\bar{z}) \\ \nu_t(z,\bar{z}) &:= (\hat{K}[[\hat{W}_t]^\dagger]\hat{K}^{-1})^{-1}(\overline{(\hat{K}[[\hat{W}_t]^\dagger]\hat{K}^{-1})})^{-1}\delta(z,\bar{z})\ . \end{aligned} \qquad \text{(II.1.1. 10)}$$

The adjoint operation is meant in the sense of $\mathcal{H}$, $\hat{K}$ means analytical extension as before and the bar means complex conjugation of the expression of the operator (i.e. any appearance of multiplication or differentiation by z is replaced with multiplication or differentiation by $\bar{z}$ and vice versa, and, of course, also numerical coefficients are complex conjugated). Here $\mu_0^{\mathbb{C}}$ and $\bar{\mu}_0^{\mathbb{C}}$ are just the analytic and anti-analytic extensions of the measure μ_0 on $\mathcal{C}$ (they are just complex conjugates of each other thanks to the positivity of μ_0) and the distribution in the second line of (II.1.1. 9) is defined by

$$\int_{\mathcal{C}_{\mathbb{C}}} d\mu_0^{\mathbb{C}}(z)d\bar{\mu}_0^{\mathbb{C}}(\bar{z})f(z,\bar{z})\delta(z,\bar{z}) = \int_{\mathcal{C}} d\mu_0(x)f(x,x) \qquad \text{(II.1.1. 11)}$$

for any smooth function f on the complexified configuration space of rapid decrease with respect to μ_0.

Whenever (II.1.1. 9) exists (it is straightforward to check that (II.1.1. 9) does the job formally), the extension of $\hat{U}_t$ to an unitary operator (isometric, densely defined and surjective) in the sense above can be expected [80]. A concrete proof is model-dependent.

In summary, we have solved *two problems in one stroke* :

We have implemented the correct adjointness relations and we have simplified the Hamiltonian (constraint) operator.

A couple of remarks are in order :

- The method does not require that $\hat{C}$ is self-adjoint, positive, bounded or at least normal. All that is important is that $\hat{W}_t$ exists on real analytic functions in the sense of Nelson's analytic vector theorem.

- It reproduces the cases of the harmonic oscillator and the case considered by Hall [79]. But it also explains *why* it works the way it works, namely it answers the question of how to identify analytic continuation with a given complex polarization of the phase space as is obvious from $\hat{K} = \hat{U}_t\hat{W}_t^{-1}$. The computation of ν_t with our metod via (II.1.1. 9), (II.1.1. 11) is considerably simpler. The harmonic oscillator corresponds to the complexifier $C = \frac{1}{2}p^2$.

- On might wonder why one should compute ν_t at all and bother with $\mathcal{H}_{\mathbb{C}}$ [82] ? Could one not just forget about the analytic continuation and work only on $\mathcal{H}$ simply by studying the spectral analysis of the unphysical operator $\hat{H}$ and defining the physical operator by $\hat{H}' := \hat{W}_t^{-1}\hat{H}\hat{W}_t$? The problem is that, while it is true that restrictions to real arguments of (generalized) eigenvectors of $\hat{H}_{\mathbb{C}}$ are *formal* eigenvectors of $\hat{H}$, these are typically not (generalized) eigenvectors in the sense of the topology of $\mathcal{H}$. Intuitively, what happens is that the measure ν_t provides for the necessary much stronger fall-off in order to turn the analytic extension of the badly behaved formal eigenvectors $\hat{W}_t^{-1}\psi$ of $\hat{H}'$ into well-defined (generalized) eigenvectors $\hat{K}\psi$ of $\hat{H}_{\mathbb{C}}$.
  One can see this also from another point of view : by unitarity, whenever $\hat{H}_{\mathbb{C}}$ is self-adjoint, so is $\hat{H}'$ but in general $\hat{H}$ is not. Thus, one would not expect the spectra of $\hat{H}, \hat{H}'$ to coincide. See the appendix of [80] for a discussion of this point.

- There are also other applications of this transform, for example in Yang-Mills theory it can be used to turn the Hamiltonian from a fourth order polynomial into a polynomial of order three only !

This completes the outline of the general framework. We will now turn to the interesting case of quantum gravity.

II.1.1.2 Wick Transform for Quantum Gravity

As Barbero [76] correctly pointed out, all the machinery that is associated with the quantum configuration space $\overline{\mathcal{A}}$ and the uniform measure μ_0 is actually also available for Lorentzian quantum general relativity if one chooses the Immirzi parameter β to be real. However, the Hamiltonian constraint then does not simplify at all as compared to the ADM expression and so the virtue of the new variables would be lost. The coherent state transform as derived below in principle combines both advantages, namely a well-defined calculus on $\overline{\mathcal{A}}$ and a simple Wheeler-DeWitt constraint.

Let us then apply the framework of the previous subsection. The phase space of Lorentzian general relativity can be given a real polarization through the canonical pair $(A_a^j := \Gamma_a^j + K_a^j, E_j^a/\kappa)$ (the case considered by Barbero with $\beta = 1$) and a complex polarization through canonical pair $(({}^{\mathbb{C}}A_a^i) := \Gamma_a^j - iK_a^j, ({}^{\mathbb{C}}E_j^a) := iE_j^a/\kappa)$ (the case considered by Ashtekar). The rescaled Hamiltonian constraint looks very simple in the complex variables, namely

$$\tilde{H}_{\mathbb{C}}(A_{\mathbb{C}}, E_{\mathbb{C}}) = \epsilon_{ijk}({}^{\mathbb{C}}F_{ab}^i)({}^{\mathbb{C}}E_j^a)({}^{\mathbb{C}}E_k^b) \qquad \text{(II.1.1. 12)}$$

but if we write $A_{\mathbb{C}}, E_{\mathbb{C}}$ in terms of A, E then the resulting Hamiltonian $\tilde{H}'(A, E)$ becomes extremely complicated. Let us compute the map W. We first of all see that we can go from (A, E) to $(A_{\mathbb{C}}, E_{\mathbb{C}})$ in a sequence of three canonical transformations given by

$$(A = \Gamma + K, E/\kappa) \to (K, E/\kappa) \to (-iK, iE/\kappa) \to (A_{\mathbb{C}} = \Gamma - iK, E_{\mathbb{C}} = iE/\kappa)\ .$$

That the first and third step are indeed canonical transformations was already shown in section I.1.3. The second step is a *phase space Wick rotation.* Since (K, E) is a canonical pair it is trivial to see that we have

$$-iK = \sum_{n=0}^{\infty} \frac{i^n}{n!}\{C, K\}_{(n)} \text{ and } iE = \sum_{n=0}^{\infty} \frac{i^n}{n!}\{C, E\}_{(n)} \qquad \text{(II.1.1. 13)}$$

where the *complexifier* or generator of the Wick transform is given by

$$C = -\frac{\pi}{2\kappa}\int_\sigma d^3x K_a^i E_i^a \qquad \text{(II.1.1. 14)}$$

which is easily seen to be the integrated densitized trace of the extrinsic curvature. C generates infinitesimal constant scale transformations. It now seems that we need to compute the generator of the transform that adds and subtracts the spin-connection Γ. However, we have seen in section I.1 that the spin-connection in three dimensions is a homogeneous polynomial of degree zero in E and its derivatives and since a *constant* scale factor is unaffected by derivatives we have $\{\Gamma, C\} = 0$. Thus in fact we have

$$A_{\mathbb{C}} = \sum_{n=0}^{\infty} \frac{i^n}{n!}\{C, A\}_{(n)} \text{ and } E_{\mathbb{C}} = \sum_{n=0}^{\infty} \frac{i^n}{n!}\{C, E\}_{(n)}\ . \qquad \text{(II.1.1. 15)}$$

The task left is to define the operator $\hat{C}$ and to compute the corresponding measure ν_t. This seems to be a very hard problem because $K_a^i = A_a^i - \Gamma_a^i$ and Γ_a^i is just a very complicated function to quantize. Nevertheless it can be done as we will see in the next section.

We conclude this section with a few remarks :

1) The Wick transform is a phase space Wick rotation and has *nothing to do with analytical continuation in the time parameter t* ! Mena Marugán [84] has given a formal relation with the usual Wick rotation corresponding to an analytical continuation of time together with a complex conformal rescaling of the four-dimensional metric.

2) As we will see in the next section, one *can* construct a well-defined operator $\hat{C}$, whether its exponential makes any sense though is an open question. But we will derive an even stronger result : one can really dispense with the complex variables altogether because one can give meaning to the *unrescaled, original* Hamiltonian constraint $H' = \tilde{H}'/\sqrt{\det(q)}$ in terms of the real variables (A, E). Although the complexifier C is then not used any more for the purpose of a Wick rotation, it still plays a crucial role in the quantization scheme displayed there. That it comes out rigorously quantized of that scheme is more a side result than a premise. The corresponding operator $\hat{H}$ which we construct directly on the Hilbert space $\mathcal{H}^0$ is surprisingly not terribly complicated. Still, it maybe important to construct a Wick transform one day because 1) it could simplify the construction of rigorous solutions and since 2) a coherent state transform always has a close connection with semiclassical physics which is important for the interpretation and the classical limit of the theory.

3) Not surprisingly, the unphysical Hamiltonian $\tilde{H}(A, E) := \tilde{H}_{\mathbb{C}}(A_{\mathbb{C}} := A, E_{\mathbb{C}} := E)$ can be recognized as the Hamiltonian constraint that one obtains from the Hamiltonian formulation of Riemannian general relativity (i.e. ordinary general relativity just that one considers four-metrics of Euclidean signature).

4) The Wick transform derived in [80] is the first honest proposal for a solution of the reality conditions for the *complex* connection variables. For a different proposal geared to a Minkowski space background, see [168].

From now on we remove the prime in H' again and will only work with physical, unrescaled functions of real variables.

II.1.2 Derivation of the Hamiltonian Constraint Operator

In view of the previous section, a crucial question that remained was whether one could make the Wick transform to work, that is, whether one could realize its generator as a self-adjoint operator on $\mathcal{H}^0$ to begin with. In the course of efforts towards this aim, a new perspective came to the foreground which enables one to get rid of the difficult complex variables altogether and to work entirely with the real ones. This then also made the existence of the Wick transform a question of marginal interest. In retrospect, it is now clear that in any case one could never have succeeded working with an operator corresponding to $\tilde{H}$ even if one could make the Wick transform work : The reason for this is so simple that it is surprising that it was not pointed out long before. It has nothing to do with the use of complex valued variables but rather with the fact that $\tilde{H}$, regardless of whether written in terms of real or complex variables, is a scalar with a density weight of *two* rather than *one*. Let us clarify this point from the outset:

When one quantizes an integrated scalar density $s(x)$ of weight k then one replaces the canonical variables by multiplication operators and functional derivatives respectively. When one applies the local operator $\hat{s}(x)$ to a state ψ, which is, in particular, a scalar of density weight zero, the various multiplication operators and functional derivatives produce a new state which is roughly of the form

$D(x)\psi'(x)$ where $\psi'(x)$ is a well-defined scalar with density weight zero and $D(x)$ is a distribution. However, the operator remembers the density weight of its classical counterpart s and therefore the density weight k must be encoded in the distribution $D(x)$. The only distribution of density weight different from zero that can appear is the delta-distribution (and derivatives thereof). We conclude that $D(x)$ is proportional to $\delta(x)^k$ (and derivatives thereof) which is meaningless unless $k=1$.

Why does one not see this problem in ordinary quantum field theory as for instance the Maxwell Hamiltonian is a density of weight two as well ? The answer is that one actually *does* see this problem : the divergence that appears can be cured in this case by normal ordering, one subtracts an infinite constant from the Hamiltonian. Such a procedure is possible in free quantum field theory on a fixed background but in background-free general relativity this cannot be done : the infinite constant contributes to the vacuum energy and cannot be discarded. Also a regularization and renormalization does not work : Consider for instance a point splitting regularization. That is, one measures distances by a background metric. If one subtracts the divergence and removes the regulator, the result is necessarily a background dependent operator destroying diffeomorphism covariance.

Actually, this problem was noted by many working on formal solutions to the Hamiltonian constraint (see, e.g., [98, 100, 99, 169, 170, 171] and references therein) but its underlying reason in terms of density weights had not been spelled out clearly.

In order to solve the problem even multiplicative renormalizations were considered, that is, one multiplies the operator by a regulator which vanishes in the limit. While this removes the background dependence one now has a quantum operator whose classical limit is zero.
Another suggestion was to take the square root of the Hamiltonian constraint $\tilde{H}$ since this reduces the density weight to one and to quantize this square root (see [172], in particular in connection with matter coupling [173]). However, since $\tilde{H}$ is famously indefinite it is unclear how to define the square root of an infinite number of non-self-adjoint, non-positive and non-commuting operators, moreover, classically the square root of a constraint has an ill-defined Hamiltonian vector field and therefore does not generate gauge transformations.
A brute force method finally to remove the singularities is to go to a lattice formulation but the problem must undoubtedly reappear when one takes the continuum limit (see, e.g., [174] and references therein).

For those reasons, *the factor* $1/\sqrt{\det(q)}$ *in* H *as compared to* $\tilde{H}$ *is, in fact, needed and one cannot work with the rescaled constraint.* Since H in either real or complex connection variables is as non-polynomial as in the ADM variables, it seems at first that the whole virtue of introducing connection variables is lost (even if the Wick transform could be made to work since the non-polynomial prefactor does not get removed by it).

However, this is by far not the case, the advantage of connection variables is twofold, there are very powerful kinematical and dynamical reasons for using them:
The kinematical reason is it has been possible to give a rigorous, background independent mathematical formulation (using real connections) only using form fields (here one forms and $(D-1)-$ forms). This has not been achieved using ADM metric variables so far. Only connections provide us with the powerful calculus on $\overline{\mathcal{A}}$.
The dynamical reason is that, as we are about to show, one can actually give rigorous meaning to H as a quantum operator on $\mathcal{H}^0$ *despite its non-polynomial nature* ! By means of a novel quantization technique the non-polynomial prefactor is absorbed into a commutator between well-defined operators. Since a commutator is essentially a derivation one can intuitively understand that this operation will express a denominator in terms of a numerator which has a better chance to be well-defined as an operator.

This technique then removed the two major roadblocks that plagued Canonical Quantum General

Relativity until 1996 in one stroke :

First, it showed that $\hat{H}$ can be made well-defined in terms of real connections and therefore, secondly, the full machinery of $L_2(\overline{\mathcal{A}}, d\mu_0)$ could be accessed.
Even more is true : the new technique turns out to be so general that it applies to any kind of field theory for which a Hamiltonian formulation exists [164, 165, 119, 184, 193, 157, 194]. The series of these papers is entitled "Quantum Spin Dynamics (QSD)" for the following reason : the Hamiltonian constraint $\hat{H}$ acting on a spin-network state creates and annihilates the spin quantum numbers with which the edges of the underlying graph are coloured. On the other hand, the ADM energy surface Hamiltonian operator [157] is essentially diagonal on spin-network states where its eigenvalue is also determined by the spin-quantum numbers. Thus, we may interpret the spin-network representation as the **non-linear Fock representation of quantum general relativity**, the spin quanta playing the role of the occupation numbers of momentum excitations of the usual Fock states of, say, Maxwell theory. The excitations of the gravitational quantum field are string-like, labelled by the edges of a graph, and the degree of freedom corresponding to an edge can be excited only according to half-integral spin quantum numbers.

The rest of this section is devoted to a hopefully pedagogical explanation of the main idea on which [164] is based. (see also [78, 175] for an even less technical introduction).

Usually, the Hamiltonian constraint is written in terms of the real connection variables as follows [76, 174] (we set $\beta = 1/2$ in this section, the generalization to arbitrary positive values is trivial, and drop the label β from all formulas)

$$H = \frac{1}{\kappa\sqrt{\det(q)}}\mathrm{tr}([F_{ab} - R_{ab}][E^a, E^b]) \tag{II.1.2. 1}$$

(we have a trace and a commutator for the Lie algebra valued quantities and kept explicitly a factor of $1/\kappa$ coming from an overall factor of $1/\kappa$ in front of the action). The reason for this clear : since A, E are the elementary variables one better avoids the appearance of $K_a^i = A_a^i - \Gamma_a^i$. We, however, will work paradoxically with the following identical formula (up to an overall numerical factor)

$$H = \frac{2}{\kappa\sqrt{\det(q)}}\mathrm{tr}([K_a, K_b][E^a, E^b]) - H_E \tag{II.1.2. 2}$$

where

$$H_E = \frac{1}{\kappa\sqrt{\det(q)}}\mathrm{tr}(F_{ab}[E^a, E^b]) \tag{II.1.2. 3}$$

is called the *Euclidean Hamiltonian constraint*, that is, the (unrescaled) unphysical Hamiltonian constraint that one would employ into the Wick rotation transform as alluded to in section II.1.1. Its natural appearance here is not a coincidence as we will see. The reason for doing this will become clear in a moment. Notice that we have correctly introduced the overall factor $1/\kappa$ in front of the action into H_E, H which will get the dimensionalities right.

Consider the following two quantities,
(i) The volume of an open region R of σ :

$$V(R) := \int_R d^3x\sqrt{|\det(q)|} \text{ and} \tag{II.1.2. 4}$$

(ii) the integrated densitized trace of the extrinsic curvature

$$K := \int_\sigma d^3x K_a^i E_i^a \qquad \text{(II.1.2. 5)}$$

the latter of which is nothing else than the generator of the Wick transform up to a factor of $-\pi/(2\kappa)$. Notice that in (II.1.2. 5) we have taken absolute values under the square root. However, $\det((q_{ab})) = [\det((e_a^i))]^2$ is anyway positive so that we could drop the absolute value at the classical level. At the quantum level, however, it will be important to keep it. On the other hand, if we define $E_j^a = \det(e_b^k)e_j^a$ then E_i^a satisfies the following anholonomic constraint

$$\det((E_i^a)) = \det((q_{ab})) \geq 0 \qquad \text{(II.1.2. 6)}$$

as pointed out before. Strictly speaking, one could argue to have to impose (II.1.2. 5) on quantum states later on. On the other hand, quantum theory is an extension of the classical theory anyway and (II.1.2. 5) could be required to hold on *semiclassical states* only in the sense of expectation values. One could also work instead with $E_j^a = |\det(e_b^k)|e_j^a$ in which case (II.1.2. 5) would no longer hold, however, then one has to absorb a factor of $\text{sgn}(\det(e_b^k))$ into the lapse function in the following formulas so that N is allowed to take both signs. Then one might want to argue that $N\text{sgn}(\det(e_b^k)) \geq 0$ should hold which is peculiar since it would mean to quantize the lapse, which is in contradiction with the whole formalism and therefore must be dropped as well in the quantum theory. Whether one strategy is preferred over the other is not yet clear. What is clear, however, is that the condition (II.1.2. 5) in the strong form (i.e. that the equality sign is excluded) is not preserved under the quantum evolution: The right hand side of (II.1.2. 5) becomes in quantum theory, roughly, the square of the volume operator. Now while the volume eigenvalues are non-negative, the value zero is attained on trivalent vertices and these are among the types of vertices created by the Hamiltonian constraint. Therefore one cannot quantize $1/\sqrt{\det(q)}$ by replacing it by the inverse volume operator. We choose here the first alternative and simply drop (II.1.2. 5) while working with $E_j^a = \det(e)e_j^a$.

The following two classical identities are *key* for all that follows :

$$\left(\frac{[E^a, E^b]_i}{\sqrt{\det(q)}}\right)(x) = \epsilon^{abc}(\text{sgn}(\det(e))e_c^i)(x) = 2\epsilon^{abc}\frac{\delta V(R)}{\delta E_i^a(x)} = 2\epsilon^{abc}\{V(R), A_a^i(x)\}/\kappa \qquad \text{(II.1.2. 7)}$$

for any region R such that $x \in R$ and

$$K_a^i(x) = \frac{\delta K}{\delta E_i^a(x)} = \{K, A_a^i(x)\}/\kappa \qquad \text{(II.1.2. 8)}$$

where (II.1.2. 8) relies on $\{\Gamma_a^i, K\} = 0$ as already pointed out in section II.1.1. In the sequel we will use the notation R_x for any open neighbourhood of $x \in \sigma$.
Using these key identities the reader can quickly convince himself that

$$\begin{aligned}(H - H_E)(x) &= -8\epsilon^{abc}\text{tr}(\{A_a, K\}\{A_b, K\}\{A_c, V(R_x)\})/\kappa^4 && \text{(II.1.2. 9)}\\ H_E(x) &= -2\epsilon^{abc}\text{tr}(F_{ab}\{A_c, V(R_x)\})/\kappa^2 && \text{(II.1.2. 10)}\end{aligned}$$

or, in integrated form, $H(N) = \int_\sigma d^3x N(x)H(x)$ etc. for some lapse function N and any smooth neighbourhood-valued function $R : x \mapsto R_x$

$$\begin{aligned}(H - H_E)(N) &= -8\int_\sigma N\text{tr}(\{A, K\} \wedge \{A, K\} \wedge \{A, V(R)\})/\kappa^4 && \text{(II.1.2. 11)}\\ H_E(N) &= -2\int_\sigma N\text{tr}(F \wedge \{A, V(R)\})/\kappa\ . && \text{(II.1.2. 12)}\end{aligned}$$

What we have achieved in (II.1.2. 9), (II.1.2. 10) or (II.1.2. 11), (II.1.2. 12) is to remove the problematic $1/\sqrt{\det(q)}$ from the denominator by means of Poisson brackets.

The reader will now ask what the advantage of all this is. The idea behind these formulas is the following :
What we want to quantize is $H(N)$ on $\mathcal{H}^0$ and since $\mathcal{H}^0$ is defined in terms of generalized holonomy variables $A(e)$ we first need to write (II.1.2. 11), (II.1.2. 12) in terms of holonomies. This can be done by introducing a triangulation $T(\epsilon)$ of σ by tetrahedra which fill all of σ and intersect each other only in lower dimensional submanifolds of σ. The small parameter ϵ is to indicate how fine the triangulation is, the limit $\epsilon \to 0$ corresponding tetrahedra of vanishing volume (the number of tetrahedra grows in this limit as to always fill out σ; we will not be specific here about what ϵ actually is, the interested reader is referred to [164]). So let $e_I(\Delta)$ denote three edges of an analytic tetrahedron $\Delta \in T(\epsilon)$ and let $v(\Delta)$ be their common intersection point with outgoing orientation (the quantities $\Delta, e_I(\Delta), v(\Delta)$, of course, also depend on ϵ but we do not display this in order not to clutter the formulae with too many symbols). The matrix consisting of the tangents of the edges $e_1(\Delta), e_2(\Delta), e_3(\Delta)$ at $v(\Delta)$ (in that sequence) has non-negative determinant which induces an orientation of Δ. Furthermore, let $a_{IJ}(\Delta)$ be the arc on the boundary of Δ connecting the endpoints of $e_I(\Delta), e_J(\Delta)$ such that the loop $\alpha_{IJ}(\Delta) = e_I(\Delta) \circ a_{IJ}(\Delta) \circ e_J(\Delta)^{-1}$ has positive orientation in the induced orientation of the boundary for $(I, J) = (1,2), (2,3), (3,1)$ and negative in the remaining cases. One can then see that in the limit as $\epsilon \to 0$ the quantities

$$
\begin{aligned}
(H^\epsilon - H^\epsilon_E)(N) &= \frac{8}{3\kappa^4} \sum_{\Delta \in T(\epsilon)} \epsilon^{IJK} N(v(\Delta)) \times \qquad \text{(II.1.2. 13)}\\
&\times \ \mathrm{tr}(h_{e_I(\Delta)}\{h^{-1}_{e_I(\Delta)}, K\} h_{e_J(\Delta)}\{h^{-1}_{e_J(\Delta)}, K\} h_{e_K(\Delta)}\{h^{-1}_{e_K(\Delta)}, V(R_{v(\Delta)})\})\\
H^\epsilon_E(N) &= \frac{2}{3\kappa^2} \sum_{\Delta \in T(\epsilon)} N(v(\Delta)) \epsilon^{IJK} \mathrm{tr}(h_{\alpha_{IJ}(\Delta)} h_{e_K(\Delta)}\{h^{-1}_{e_K(\Delta)}, V(R_{v(\Delta)})\})
\end{aligned}
$$

(II.1.2. 14)

converge to (II.1.2. 11), (II.1.2. 12) respectively pointwise on $\mathcal{M}$ *for any choice of triangulation* ! This independence of the limit, for the classical theory, from the choice of the family of triangulations enables us to choose the triangulations state-dependent just as for the area operator, see below.

Suppose now that we can turn $V(R)$ and K into well-defined operators on $\mathcal{H}$, densely defined on cylindrical functions. Then, according to the rule that upon quantization one should replace Poisson brackets by commutators times $1/(i\hbar)$ (II.1.2. 13), (II.1.2. 14) *would become densely defined regulated operators on $\mathcal{H}^0$ without any divergences for a specific choice of factor ordering* ! We will discuss the issue of what happens upon removal of the regulator ϵ in a moment.

Is it then true that $\hat{V}(R)$ and $\hat{K}$ exist ? We have seen in section I.4 that the answer is affirmative for the case of the volume operator. We use the version of the volume operator that was constructed in [153] as compared to the one in [150] because it turns out that only the operator [153] gives a densely defined Hamiltonian constraint operator in the regularization scheme that we advertize here, it is important that the volume vanishes on planar vertices (that is, the tangent space at the vertex spanned there by the tangents of the edges incident at it is at most two-dimensional) .

Recall from section I.4 that the volume operator of [153] acts on a function cylindrical over a graph γ as follows :

$$\hat{V}(R) f_\gamma := \frac{\ell_p^3}{4} \sum_{v \in V(\gamma) \cap R} \sqrt{\left| \frac{i}{3!} \sum_{e \cap e' \cap \tilde{e} = v} \epsilon(e, e', \tilde{e}) \epsilon_{ijk} R^i_e R^j_{e'} R^k_{\tilde{e}} \right|}\, f_\gamma \qquad \text{(II.1.2. 15)}$$

where the sum is over the set $V(\gamma)$ of all vertices v of the graph γ that lie in R and over all unordered triples of edges that start at v (we can take the orientation of each edge incident at v to be outgoing by suitably splitting an edge into two halves if necessary). The function $\epsilon(e, e', \tilde{e})$ takes the values $+1, -1, 0$ if the tangents of the three edges at v (in that sequence) form a matrix of positive, negative or vanishing determinant and the right invariant vector fields R^i_e were defined in section I.3. The absolute value $|\hat{B}|$ of the operator $\hat{B}$ indicates that one is supposed to take the square root of the operator $\hat{B}^\dagger \hat{B}$. The dense domain of this operator are the thrice differentiable cylindrical functions. Notice that planar vertices of arbitrary valence do not contribute. Surprisingly, also arbitrary tri-valent vertices do not contribute [151] if the corresponding state is gauge-invariant. (Proof: We have $-(R^j_1 + R^j_2) = R^j_3$ due to gauge invariance where $R^j_I = R^j_{e_I}$, $I = 1, 2, 3$. Substituting this into $\epsilon_{jkl} R^j_1 R^j_2 R^j_3$ and using $[R^j_I, R^k_J] = -2\delta_{IJ}\epsilon_{jkl} R^l_I$ completes the proof).

Thus, it seems that one can make sense out of a regulated operator corresponding to (II.1.2. 13) for each N, in particular for $N = 1$. Now recall the classical identity that the integrated densitized trace of the extrinsic curvature is the "time derivative" of the total volume

$$K = \{H^E(1), V(\sigma)\} = \{H(1), V(\sigma)\} . \tag{II.1.2. 16}$$

where $N = 1$ is the constant lapse equal to unity. This formula makes sense even if σ is not compact (see [164] for the details). Notice that (II.1.2. 16) holds for either signature (i.e. it does not matter which Hamiltonian constraint of the Hamiltonian formulation of general relativity we use, the one corresponding to four metrics of Euclidean or Lorentzian signature). But if we then replace again Poisson brackets by commutators times $1/(i\hbar)$ and define

$$\hat{K}^\epsilon := -\frac{i}{\hbar}[\hat{H}^\epsilon_E(1), \hat{V}(\sigma)] \tag{II.1.2. 17}$$

using the already defined quantities $\hat{H}^\epsilon(1), \hat{V}(\sigma)$ it seems that we can also define a regulated operator corresponding to (II.1.2. 14) !

This concludes the explanation of the main idea. The next subsection comments on the concrete implementation of this idea.

II.1.3 Mathematical Definition of the Hamiltonian Constraint Operator

Obviously, central questions regarding the concrete implementation of the technique are :
I) What are the allowed, physically relevant choices for a family of triangulations $T(\epsilon)$?
II) How should one treat the limit $\epsilon \to 0$ for the operator $\hat{H}^\epsilon(N)$? That is, should one keep ϵ finite and just refine $\gamma \to \sigma$ or is there an operator topology such that this limit can be given a meaning ? Secondly, does the refined or limit operator remember something about the choice of the family $T(\epsilon)$ or is there some notion of universality ?
III) What is the commutator algebra of these (limits of) operators, is it free of anomalies ?
We will address these issues separately.

II.1.3.1 Concrete Implementation

A natural choice for a triangulation turns out to be the following (we simplify the presentation drastically, the details can be found in [164]):
Given a graph γ one constructs a triangulation $T(\gamma, \epsilon)$ of σ *adapted* to γ which satisfies the following basic requirements :

a) The graph γ is embedded in $T(\gamma,\epsilon)$ for all $\epsilon > 0$.

b) The valence of each vertex v of γ, viewed as a vertex of the infinite graph $T(\epsilon,\gamma)$, remains constant and is equal to the valence of v, viewed as a vertex of γ, for each $\epsilon > 0$.

c) Choose a system of analytic arcs $a^\epsilon_{\gamma,v,e,e'}$, one for each pair of edges e, e' of γ incident at a vertex v of γ, which do not intersect γ except in its endpoints where they intersect transversally. These endpoints are interior points of e, e' and are those vertices of $T(\epsilon,\gamma)$ contained in e, e' closest to v for each $\epsilon > 0$ (i.e. no others are in between). For each $\epsilon, \epsilon' > 0$ the arcs $a^\epsilon_{\gamma,v,ee'}, a^{\epsilon'}_{\gamma,v,e,e'}$ are diffeomorphic with respect to analytic diffeomorphisms. The segments of e, e' incident at v with outgoing orientation that are determined by the endpoints of the arc $a^\epsilon_{\gamma,v,e,e'}$ will be denoted by $s^\epsilon_{\gamma,v,e}, s^\epsilon_{\gamma,v,e'}$ respectively. Finally, if φ is an analytic diffeomorphism then $s^\epsilon_{\varphi(\gamma),\varphi(v),\varphi(e)}, a^\epsilon_{\varphi(\gamma),\varphi(v),\varphi(e),\varphi(e')}$ and $\varphi(s^\epsilon_{\gamma,v,e}), \varphi(a^\epsilon_{\gamma,v,e,e'})$ are analytically diffeomorphic.

d) Choose a system of mutually disjoint neighbourhoods $U^\epsilon_{\gamma,v}$, one for each vertex v of γ, and require that for each $\epsilon > 0$ the $a^\epsilon_{\gamma,v,e,e'}$ are contained in $U^\epsilon_{\gamma,v}$. These neighbourhoods are nested in the sense that $U^\epsilon_{\gamma,v} \subset U^{\epsilon'}_{\gamma,v}$ if $\epsilon < \epsilon'$ and $\lim_{\epsilon\to 0} U^\epsilon_{\gamma,v} = \{v\}$.

e) Triangulate $U^\epsilon_{\gamma,v}$ by tetrahedra $\Delta(\gamma, v, e, e', \tilde{e})$, one for each ordered triple of distinct edges $e, e', \tilde{e}$ incident at v, bounded by the segments $s^\epsilon_{\gamma,v,e}, s^\epsilon_{\gamma,v,e'}, s^\epsilon_{\gamma,v,\tilde{e}}$ and the arcs $a^\epsilon_{\gamma,v,e,e'}, a^\epsilon_{\gamma,v,e',\tilde{e}}, a^\epsilon_{\gamma,v,\tilde{e},e}$ from which loops $\alpha^\epsilon(\gamma; v; e, e')$ etc. are built and triangulate the rest of σ arbitrarily. The ordered triple $e, e', \tilde{e}$ is such that their tangents at v, in this sequence, form a matrix of positive determinant.

Requirement a) prevents the action of the Hamiltonian constraint operator from being trivial. Requirement b) guarantees that the regulated operator $\hat{H}^\epsilon(N)$ is densely defined for each ϵ. Requirements c), d) and e) specify the triangulation in the neighbourhood of each vertex of γ and leave it unspecified outside of them. The more detailed prescription of [164] shows that triangulations satisfying all of these requirements always exist and can also deal with degenerate situations, e.g., how to construct a tetrahedron for a planar vertex. More specifically, what we have done in [164] is to fix the routing of the analytical arcs through the "forest" of the already present edges in such a way that it is invariant under smooth diffeomorphisms that leave γ invariant and the arcs analytic. The use of smooth diffeomorphisms here is not in contradiction to having only an analytic manifold as we use them only in order to choose the routing, they do not play any other role e.g. in imposing the diffeomorphism constraint. In particular, since we do not need that arcs corresponding to different pairs of edges are analytically diffeomorphic, there is no contradiction. Here we are more general than in [164] in that we just use the *axiom of choice*. That is, we only use that a choice function

$$a^\epsilon : \Gamma^\omega_0 \to \Gamma^\omega_0 \; \gamma \mapsto \{a^\epsilon_{\gamma,v,e,e'}\}_{v\in V(\gamma);\, e,e'\in E(\gamma);\, v\in\partial e\cap\partial e'} \tag{II.1.3. 1}$$

subject to requirements a) – e) always exists.

The reason for why those tetrahedra lying outside the neighbourhoods of the vertices described above are irrelevant rest crucially on the choice of ordering (II.1.2. 14) with $[\hat{h}_s^{-1}, \hat{V}]$ on the most right and on our choice of the volume operator [153]: If f is a cylindrical function over γ and s has support outside the neighbourhood of any vertex of γ, then $V(\gamma \cup s) - V(\gamma)$ consists of planar at most four-valent vertices only so that $[\hat{h}_s^{-1}, \hat{V}]f = 0$. Notice, however, that [150] does not vanish on planar vertices and so $[\hat{h}_s^{-1}, \hat{V}]f$ would not vanish even on trivalent vertices in $V(\gamma \cup s) - V(\gamma)$ because it is not gauge invariant. In other words, in the limit of small ϵ the operator would map us out of the space of cylindrical functions. Therefore the Hamiltonian constraint operator inherits from the volume operator a basic property : It annihilates all states cylindrical with respect to graphs with only co-planar vertices as can be understood from the fact that the volume operator enters the construction of both $\hat{H}^\epsilon_E(N), \hat{H}^\epsilon(N)$. In other words, the dynamics "happens only at the vertices of a graph".

Notice that a)-e) are natural extensions to arbitrary graphs of what one does in lattice gauge theory [176] with one exception : what we will get is not an operator $\hat{H}^\epsilon(N)$ to begin with, but actually a family of operators $\hat{H}^\epsilon_\gamma(N)$, one for each graph γ. This happened because we adapted the triangulation to the graph of the state on which the operator acts. One must then worry that this does not define a linear operator any more, that is, that it is not cylindrically consistently defined. Here we circumvent that problem as follows: We do not define the operator on functions cylindrical over graphs but cylindrical over *coloured graphs*, that is, we define it on spin-network functions. Since every smooth cylindrical function, the domain for the operator that we will choose, is a finite linear combination of spin-network functions this defines the operator uniquely as a linear operator. Any operator automatically becomes consistent if one defines it on a basis, the consistency condition simply drops out.

Moreover, the regulated operator $\hat{H}^\epsilon(N)$ is by construction background independently defined for each ϵ but not symmetric which, as described in section III.7, is not a necessary requirement for a constraint operator and even argued to be better not the case [177] in order for the constraint algebra to be non-anomalous for open constraint algebras.

Finally, we point out that beyond the freedom of a choice function (II.1.3. 1) even requirements a)-e) can be generalized and even the regularization itself can be generalized. For instance in [178] one uses instead of $\mathrm{tr}(\tau_j h_\alpha)$ the function

$$\frac{\sum_{k=1}^N \mathrm{tr}(\tau_j h_\alpha^{n_k})}{\sum_{k=1}^N n_k} \tag{II.1.3. 2}$$

for any choice of integers n_k such that the denominator is non-vanishing which again gives the correct continuum limit since all the functions (II.1.3. 1) are identical in the leading order that we need. Hence, there is vast room for generalizations. Which choice is "more physical" than another, whether they all are equivalent or whether all of them are unphysical can only be decided in the investigation of the classical limit.

Let us then display the action of the Hamiltonian constraint on a spin-network function f_γ cylindrical with respect to a graph γ. It is given by

$$\hat{H}^\epsilon_E(N) f_\gamma = \frac{16}{3i\kappa\ell_p^2} \sum_{v\in V(\gamma)} \frac{N(v)}{E(v)} \sum_{v(\Delta)=v} \epsilon^{IJK} \mathrm{tr}(h_{\alpha_{IJ}(\Delta)} h_{e_K(\Delta)} [h^{-1}_{e_K(\Delta)}, \hat{V}(U_\epsilon(v))])\, f_\gamma \tag{II.1.3. 3}$$

$$\begin{aligned}(\hat{H}^\epsilon - \hat{H}^\epsilon_E)(N) f_\gamma &= \frac{64}{3\kappa(i\ell_p^2)^3} \sum_{v\in V(\gamma)} \frac{N(v)}{E(v)} \sum_{v(\Delta)=v} \epsilon^{IJK} \times \\ &\times\ \mathrm{tr}(h_{e_I(\Delta)}[h^{-1}_{e_I(\Delta)}, \hat{K}^\epsilon] h_{e_J(\Delta)} [h^{-1}_{e_J(\Delta)}, \hat{K}^\epsilon] h_{e_K(\Delta)} [h^{-1}_{e_K(\Delta)}, \hat{V}(U_\epsilon(v))])\, f_\gamma\end{aligned} \tag{II.1.3. 4}$$

where $\hat{K}_\epsilon$ is defined by (II.1.2. 17). The first sum is over all the vertices of a graph and the second sum over all ordered tetrahedra of the triangulation $T(\epsilon, \gamma)$ that saturate the vertex (the remaining tetrahedra drop out). The symbols $e_I(\Delta)$ etc. mean the same as in (II.1.2. 13), (II.1.2. 14) just that now the tetrahedra in question are the particular ones as specified in a)-e) above. Here the numerical factors $E(v) = \left(\begin{array}{c} n(v) \\ 3 \end{array} \right)$, where $n(v)$ is the valence of the vertex v, come about as follows:

Given a triple of edges (e, e', e'') incident at v with outgoing orientation consider the tetrahedron $\Delta^\epsilon(\gamma, v, e, e', e'')$ bounded by the three segments $s^\epsilon_{\gamma,v,e} \subset e, s^\epsilon_{\gamma,v,e'} \subset e', s^\epsilon_{\gamma,v,e''} \subset e''$ incident at v and the three arcs $a^\epsilon_{\gamma,v,e,e'}, a^\epsilon_{\gamma,v,e',e''}, a^\epsilon_{\gamma,v,e'',e}$. We now define the "mirror images"

$$\begin{aligned} s^\epsilon_{\gamma,v,\bar{p}}(t) &:= 2v - s^\epsilon_{\gamma,v,p}(t) \\ a^\epsilon_{\gamma,v,\bar{p},\bar{p}'}(t) &:= 2v - a^\epsilon_{\gamma,v,p,p'}(t) \\ a^\epsilon_{\gamma,v,\bar{p},p'}(t) &:= a^\epsilon_{\gamma,v,\bar{p},\bar{p}'}(t) - 2t[v - s^\epsilon_{\gamma,v,p'}(1)] \\ a^\epsilon_{\gamma,v,p,\bar{p}'}(t) &:= a^\epsilon_{\gamma,v,p,p'}(t) + 2t[v - s^\epsilon_{\gamma,v,p'}(1)] \end{aligned} \tag{II.1.3. 5}$$

where $p \neq p' \in \{e, e', e''\}$ and we have chosen any parameterization of segments and arcs. Using the data (II.1.3. 5) we build seven more "virtual" tetrahedra bounded by these quantities so that we obtain altogether eight tetrahedra that saturate v and triangulate a neighbourhood $U^\epsilon_{\gamma,v,e,e',\tilde{e}}$ of v. Let $U^\epsilon_{\gamma,v}$ be the union of these neigbourhoods as we vary the ordered triple of edges of γ incident at v. The $U^\epsilon_{\gamma,v}$, $v \in V(\gamma)$ were chosen to be mutually disjoint in point d) above. Let now

$$\begin{aligned} \bar{U}^\epsilon_{\gamma,v,e,e',e''} &:= U^\epsilon_{\gamma,v} - U^\epsilon_{\gamma,v,e,e',e''} \\ \bar{U}^\epsilon_\gamma &:= \sigma - \bigcup_{v \in V(\gamma)} U^\epsilon_{\gamma,v} \end{aligned} \tag{II.1.3. 6}$$

then we may write any classical integral (symbolically) as

$$\begin{aligned} \int_\sigma &= \int_{\bar{U}^\epsilon_\gamma} + \sum_{v \in V(\gamma)} \int_{U^\epsilon_{\gamma,v}} \\ &= \int_{\bar{U}^\epsilon_\gamma} + \sum_{v \in V(\gamma)} \frac{1}{E(v)} \sum_{v=b(e)\cap b(e')\cap b(e'')} [\int_{U^\epsilon_{\gamma,v,e,e',e''}} + \int_{\bar{U}^\epsilon_{\gamma,v,e,e',e''}}] \\ &\approx \int_{\bar{U}^\epsilon_\gamma} + \sum_{v \in V(\gamma)} \frac{1}{E(v)} [\sum_{v=b(e)\cap b(e')\cap b(e'')} 8 \int_{\Delta^\epsilon_{\gamma,v,e,e',e''}} + \int_{\bar{U}^\epsilon_{\gamma,v,e,e',e''}}] \end{aligned} \tag{II.1.3. 7}$$

where in the last step we have noticed that classically the integral over $U^\epsilon_{\gamma,v,e,e',e''}$ converges to eight times the integral over $\Delta^\epsilon_{\gamma,v,e,e',e''}$. Now when triangulating the regions of the integrals over $\bar{U}_{v,e,e',e''}$ in (II.1.3. 7), regulation and quantization gives operators that vanish on f_γ because the corresponding regions do not contain a non-planar vertex of γ.

Notice that (II.1.3. 3) and (II.1.3. 4) are *finite* for each $\epsilon > 0$, that is, densely defined without that any renormalization is necessary and with range in the smooth cylindrical functions again. Furthermore, the adjoints of the expressions (II.1.3. 3) and (II.1.3. 4) are densely defined on smooth cylindrical functions again so that we get in fact a consistently and densely defined family of closed operators on $\mathcal{H}^0$.

Let us check the dimensionalities: The volume operator in (II.1.3. 3) is given by ℓ_p^3 times a dimension free operator, hence (II.1.3. 3) is given by $\ell_p/\kappa = m_p$ times a dimension free operator. Hence the correct dimension of Planck mass $m_p = \sqrt{\hbar/\kappa}$ has popped out. Therefore, by inspection, (II.1.2. 17) has dimension of $\ell_p^3 m_p/\hbar = \ell_p^2$ which is correct since $K(x) = \sqrt{\det(q)}(x) K_{ab}(x) q^{ab}(x)$ dimension cm^{-1} so that $K = \int d^3x K(x)$ has dimension cm^2. Finally therefore (II.1.3. 4) has the correct dimension of $(\ell_p^2)^2 \ell_p^3/(\kappa \ell_p^6 = m_p$ again.

II.1.3.2 Operator Limits

Basically there are two, technically equivalent viewpoints towards treating the limit $\epsilon \to 0$.

A) *Effective Operator Viewpoint*
The more radical proposal is *to drop the parameter* ϵ from all formulas. That is, take a choice function a once and for all. One gets a densely defined family of closed operators. One may object that on a given graph γ one does not get a quantization of the full expressions (II.1.2. 13), (II.1.2. 14), however, that is only because the graph γ does not fill all of σ. In other words, the continuum limit of infinitely fine triangulation of the Riemann sum expressions (II.1.2. 13), (II.1.2. 14) in the classical theory is nothing else than taking the graphs, on which the operator is probed, finer and finer. This is a new viewpoint not previously reported in the literature and could be called the *effective operator viewpoint* because on fine but not infinitely fine graphs the classical limit of the operator will only *approximate* the exact classical expression in the same way as (II.1.2. 13) and (II.1.2. 14) only approximate (II.1.2. 11) and (II.1.2. 12). However, it may be that this *is the fundamental theory* and classical physics is just an approximation to it. This way the UV regulator ϵ corresponding to the continuum limit is trivially removed and our family of operators is really defined on $\mathcal{H}^0$. Whether the operator $\hat{H}$ that we then obtain has the correct classical limit cannot decided at this stage but is again subject to a rigorous semiclassical analysis which requires new input, see section II.3.

B) *Limit Operator Viewpoint*
The challenge is to find an operator topology in which the one-parameter family of operators $\hat{H}^\epsilon$ converges.
The operators (II.1.3. 3) and (II.1.3. 4) are easily seen to be unbounded (already the volume operator has this property). Thus, a convergence in the uniform or strong operator topology is excluded. Next, one may try the weak operator topology (matrix elements converge pointwise) but with respect to this topology the limit would be the zero operator (it is too weak) : for instance, a matrix element between two spin-network states is non-zero for at most one value of ϵ. Finally, we try the weak* topology, that is, we must check whether $\Psi(\hat{H}^\epsilon(N)f)$ converges for each $\Psi \in \mathcal{D}', f \in \mathcal{D}$ where $\mathcal{D} = C^\infty(\overline{\mathcal{A}})$ with its natural nuclear topology is a dense domain and $\mathcal{D}'$ is its topological dual. It turns out that this topology is a little bit too strong, however, convergence holds with respect to a topology which we might call *Uniform Rovelli-Smolin Topology* (URST) in appreciation of the fact that Rovelli and Smolin first pointed out in [172] that, if instead of $\mathcal{D}'$ we consider the space $\mathcal{D}^*$ of *diffeomorphism invariant algebraic distributions* on $\mathcal{D}$, then objects of the form $\Psi(\hat{H}^\epsilon(N)f)$ do not depend at all on the position or shape of the arcs $a^\epsilon_{\gamma,v,e,e'}$ alluded to above. In their original work [172] Rovelli and Smolin did not spell out this property in the context of $\mathcal{H}^0$ and also they did not have a well-defined constraint operator but their observation applies to a huge class of operators, their only feature being an analog of property c) above. This is how we proceeded in [164, 165, 119].

Therefore, since all the triangulations $T(\gamma,\epsilon)$ restricted to each of the neighbourhoods $U^\epsilon_{\gamma,v}$ are diffeomorphic by property c) above, the numbers $\Psi(\hat{H}^\epsilon(N)f)$ are actually *already independent of* ϵ ! Therefore we have the striking result that with respect to the URST

$$\hat{H}(N) := \lim_{\epsilon\to 0} \hat{H}^\epsilon(N) = \hat{H}^{\epsilon_0}(N) \tag{II.1.3. 8}$$

where ϵ_0 is an arbitrary but fixed positive number. Notice that we require that for each $\delta > 0$ there exists an $\epsilon'(\delta) > 0$ such that for each $f \in \mathcal{D}, \Psi \in \mathcal{D}^*_{Diff}$

$$|\Psi(\hat{H}^\epsilon(N)f) - \Psi(\hat{H}^{\epsilon_0}(N)f)| < \delta$$

for all $\epsilon < \epsilon'(\delta)$ where $\epsilon'(\delta)$ depends only on δ but not on f, Ψ. In other words, we have convergence *uniform* in $\mathcal{D} \times \mathcal{D}^*_{Diff}$ rather than pointwise. This will be important in what follows.

Notice that therefore the convergence in the URST is very similar to the effective operator viewpoint in the sense that it gives a topology in which it is allowed to drop the label ϵ from the choice function altogether.

In particular we stress that *in contrast to the viewpoint taken in [179, 180] we still have the operator defined on $\mathcal{H}^0$ and not on the dual subspace $\mathcal{D}^*_{Diff} \subset \mathcal{D}^*$ or an extension thereof,* precisely in the same sense as the limit of a family of operators which converges in the weak * topology on $\mathcal{D}$ is still considered an operator on $\mathcal{D}$ and not a dual operator on $\mathcal{D}'$. In fact, the dual of $\hat{H}(N)$ cannot be defined on $\mathcal{D}^*_{Diff}$ because that space is not left invariant by $\hat{H}(N)'$ as we pointed out frequently which is why the authors of [179, 180] have to take an extension to the so-called "vertex smooth" distributions $\mathcal{D}^*_{Diff} \subset \mathcal{D}^*_\star \subset \mathcal{D}^*$ which is genuinely bigger than $\mathcal{D}^*_{Diff}$ and therefore unphysical. Our viewpoint is completely different: *We do not want to define $\hat{H}'(N)$ at all, we just use $\mathcal{D}^*_{Diff}$ as a means to define a topology*!

On the other hand, the physical reason for why testing convergence of the operator only on $\mathcal{D}^*_{Diff}$ rather than on a bigger space is precisely because we are eventually going to look for the space of solutions to all constraints which in turn must be a subspace $\mathcal{D}^*_{phys}$ of $\mathcal{D}^*_{Diff}$, so in a sense we do not need stronger convergence. Notice that $\mathcal{D}^*_{phys}$ *is* left invariant by the dual action of $\hat{H}(N)$ (namely it is mapped to zero).

Again, whether the continuum operator thereby obtained has the correct classical limit must be decided in an additional step.

Which viewpoint one takes is a matter of taste, technically they are completely equivalent. The limit operator viewpoint has the advantage that it shows that many choice functions are going to be physically equivalent and thus decreases (but does not remove) the degree of redundancy. In what follows we will therefore drop the label ϵ.

The limit (II.1.3. 8) certainly only depends on the diffeomorphism invariant characteristics of the particular triangulation $T(\gamma, \epsilon)$ that we chose. For instance, the limit would be different if we would use arcs that intersect the graph tangentially or which are smooth rather than analytical. Other than that, there is no residual "memory" of the triangulation.

II.1.3.3 Commutator Algebra

We now come to question III) whether the commutator between two Hamiltonian constraints and between Hamiltonian and diffeomorphism constraints exists and is free of anomalies.

1) *Hamiltonian and Diffeomorphism Constraint*
Recall that the infinitesimal generator of diffeomorphisms is ill-defined so that we must check the commutator algebra in terms of finite diffeomorphisms. The classical infinitesimal relation $\{\vec{H}(u), H(N)\} = H(u[N])$ can be exponentiated and gives

$$e^{t\mathcal{L}_{\chi_{\vec{H}(u)}}} \cdot H(N) = H((\varphi^u_t)^* N)$$

where $\chi_{\vec{H}(u)}$ denotes the Hamiltonian vector field of $\vec{H}(u)$ on the classical continuum phase space $\mathcal{M}$ and φ^u_t the one parameter family of diffeomorphisms generated by the integral curves of the vector field u. It tells us that $H(x)$ is a scalar density of weight one. Therefore we expect to have in quantum theory the relation

$$\hat{U}(\varphi)\hat{H}(N)\hat{U}(\varphi)^{-1} = \hat{H}(\varphi^* N), \tag{II.1.3. 9}$$

To check whether (II.1.3. 9) is satisfied, we notice that for a spin network function f_γ we have by the definition of the action of the diffeomorphism group $\hat{U}(\varphi)$ on $\mathcal{H}^0$ on the one hand

$$\begin{aligned}
\hat{U}(\varphi)\hat{H}(N)f_\gamma &= \hat{U}(\varphi)\sum_{v\in V(\gamma)} N(v)\hat{H}_{v,a(\gamma)}f_\gamma \\
&= \sum_{v\in V(\gamma)} N(v)\hat{H}_{\varphi^{-1}(v),\varphi^{-1}(a(\gamma))}f_{\varphi^{-1}(\gamma)} \\
&= \sum_{v\in V(\gamma)} (\varphi^* N)(\varphi^{-1}(v))\hat{H}_{\varphi^{-1}(v),\varphi^{-1}(a(\gamma))}f_{\varphi^{-1}(\gamma)} \\
&= [\hat{U}(\varphi)\hat{H}(N)\hat{U}(\varphi)^{-1}]\hat{U}(\varphi)f_\gamma \\
&= [\hat{U}(\varphi)\hat{H}(N)\hat{U}(\varphi)^{-1}]f_{\varphi^{-1}(\gamma)}
\end{aligned} \tag{II.1.3. 10}$$

and on the other hand

$$\begin{aligned}
\hat{H}(\varphi^* N)f_{\varphi^{-1}(\gamma)} &= \sum_{v\in V(\varphi^{-1}(\gamma))} (\varphi^* N)(v)\hat{H}_{v,a(\varphi^{-1}(\gamma))}f_{\varphi^{-1}(\gamma)} \\
&= \sum_{v\in V(\gamma)} (\varphi^* N)(\varphi^{-1}(v))\hat{H}_{\varphi^{-1}(v),a(\varphi^{-1}(\gamma))}f_{\varphi^{-1}(\gamma)}
\end{aligned} \tag{II.1.3. 11}$$

Here $\hat{H}_{v,a(\gamma)}$ is the operator coefficient of $N(v)$ in (II.1.3. 3), (II.1.3. 4) which depends on the graph $a(\gamma)$ assigned to γ through the choice function a, that is, the segments $s_{\gamma,v,e}$ and arcs $a_{\gamma,v,e,e'}$. Comparing (II.1.3. 10) and (II.1.3. 11) we get equality provided that

$$\varphi\circ a = a\circ\varphi \;\forall\varphi\in \mathrm{Diff}^\omega(\sigma) \tag{II.1.3. 12}$$

This seems to burden us with the proof that such a choice function really exists and in fact we do not have a proof although it would be very nice to have one since it would decrease the possible number of choice functions. However, we can avoid this by the observation that our choice function was constructed in such a way that the assignments $a(\gamma)$ and $a(\varphi(\gamma))$ are analytically diffeomorphic. In other words we always find an analytical diffeomorphism $\varphi'_{\varphi^{-1}(\gamma)}$ which preserves $\varphi^{-1}(\gamma)$ such that

$$[\hat{U}(\varphi)\hat{H}(N)\hat{U}(\varphi)^{-1}]f_{\varphi^{-1}(\gamma)}[\hat{U}(\varphi'_{\varphi^{-1}(\gamma)})\hat{H}(\varphi^* N)\hat{U}(\varphi'_{\varphi^{-1}(\gamma)})^{-1}]f_{\varphi^{-1}(\gamma)} \tag{II.1.3. 13}$$

for any γ and any $f_{\varphi^{-1}(\gamma)}$. Thus, while (II.1.3. 9) is violated, it is violated in an allowed way because the "anomaly" is a constraint operator again. Put differently, the "anomaly" is not seen in the URST so that (II.1.3. 9) is an *exact operator identity* in the URST.

In that sense then, $\hat{H}(N)$ is a diffeomorphism covariant, densely defined, closed operator on $\mathcal{H}^0$.

2) *Hamiltonian and Hamiltonian Constraint*
There are three important properties of the operator $\hat{H}(N)$ that follow from our class of choice functions (properties a)-e)) :

A) First of all, we observe that $\hat{H}(N)$ has dense domain and range consisting of smooth (in the sense of $\mathcal{D}$) cylindrical functions. Therefore it makes sense to multiply operators and in particular to compute commutators.
B) Secondly, it annihilates planar vertices.
C) Thirdly, *for no other choice of triangulation proposed so far other than the one we proposed in [164] and only when using the volume operator of [153] rather than the one of [150]* is it true that in fact any finite product of operators $\hat{H}^{\epsilon_1}(N_1)..\hat{H}^{\epsilon_n}(N_n)$ is independent of the parameters $\epsilon_1,..,\epsilon_n$ in the URST.

The second and third properties do not hold for a more general class of operators considered in the papers [179, 180] so that there is no convergence in the URST not even of the operators themselves, not to speak of their commutators. Since certainly none (of the duals) of these operators leaves the space $\mathcal{D}^*_{Diff}$ invariant, in order to compute commutators these authors suggest to introduce the larger, unphysical space $\mathcal{D}^*_\star$ already mentioned on which one can compute limits $\hat{H}'(N)=\lim_{\epsilon\to 0}(\hat{H}^\epsilon(N))'$ *pointwise* in $\mathcal{D}^*_\star\times\mathcal{D}$ of their duals and products of these limits.

Let again f_γ be a spin-network function over some graph γ. Then we compute

$$[\hat{H}(N),\hat{H}(N')]f_\gamma = \sum_{v\in V(\gamma)}[N'(v)\hat{H}(N)-N(v)\hat{H}(N')]\hat{H}_{a(\gamma)_{|v}}f_\gamma \tag{II.1.3. 14}$$
$$= \sum_{v\in V(\gamma)}\sum_{v'\in V(\gamma)\cup a(\gamma)_{|v})}[N'(v)N(v')-N(v)N'(v')]\hat{H}_{a(\gamma\cup a(\gamma)_{|v})_{|v'}}\hat{H}_{a(\gamma)_{|v}}f_\gamma$$

where for clarity we have written $\hat{H}_{a(\gamma)_{|v}}\equiv\hat{H}_{v,a(\gamma)}$ in order to indicate that $\hat{H}_{v,a(\gamma)}$ does not depend on all of $a(\gamma)$ but only on its restriction to the arcs and segments around v. We are abusing somewhat the notation in the second step because one should really expand $\hat{H}_{a(\gamma)_{|v}}f_\gamma$ into spin network functions over $\gamma\cup a(\gamma)_{|v}$ and then apply the second operator to that expansion into spin-network functions. In particular, $\hat{H}_{a(\gamma)_{|v}}f_\gamma$ it really is a finite linear combination of terms where each of them depends only on $\gamma\cup a(\gamma)_{|v,e,e'}$ for some edges e,e' incident at v and each of those should be expanded into spin-network functions. We will not write this explicitly because it is just a book keeping exercise and does not change anything in the final argument. So either one writes out all the details or one just assumes for the sake of the argument that $\hat{H}_{a(\gamma)_{|v}}f_\gamma$ is a spin network function over $\gamma\cup a(\gamma)_{|v}$. Everything we say is more or less obvious for the Euclidean Hamiltonian constraint but a careful analysis shows that it extends to the Lorentzian one as well.

Let us now analyze (II.1.3. 14). The right hand side surely vanishes for $v'=v$. We notice that any vertex $v'\in V(\gamma\cup a(\gamma)_{|v})-V(\gamma)$ is planar and since $\hat{H}_{v',a(\gamma\cup a(\gamma))}$ has an operator of the form $[h_s^{-1},\hat{V}]$ to the outmost right hand side where s is a segment, incident at v', of an edge incident at v', it follows that none of these vertices contributes. Here it was again crucial that we used the operator [153] rather than the operator [150] ! Thus (II.1.3. 14) reduces to

$$\begin{aligned}[\hat{H}(N),\hat{H}(N')]f_\gamma &= \sum_{v\neq v'\in V(\gamma)}[N'(v)N(v')-N(v)N'(v')]\hat{H}_{a(\gamma\cup a(\gamma)_{|v}))_{|v'}}\hat{H}_{a(\gamma)_{|v}}f_\gamma\\ &= \frac{1}{2}\sum_{v\neq v'\in V(\gamma)}[N'(v)N(v')-N(v)N'(v')]\times\\ &\quad\times[\hat{H}_{a(\gamma\cup a(\gamma)_{|v}))_{|v'}}\hat{H}_{a(\gamma)_{|v}}-\hat{H}_{a(\gamma\cup a(\gamma)_{|v'}))_{|v}}\hat{H}_{a(\gamma)_{|v'}}]f_\gamma\end{aligned} \tag{II.1.3. 15}$$

where in the second step we used the antisymmetry of the expression $[N'(v)N(v')-N(v)N'(v')]$ in v, v'. Now the crucial point is that for $v \neq v' \in V(\gamma)$ the prescription of how to attach the arcs first around v and then around v' as compared to the opposite may not be the same because our prescription depends explicitly on the graph to which we apply it, however, they are certainly analytically diffeomorphic.

Thus, there exist analytical diffeomorphisms $\varphi_{\gamma,v,v'}$ preserving $\gamma \cup a(\gamma)_{|v}$ such that

$$\hat{H}_{a(\gamma\cup a(\gamma)_{|v}))_{|v'}}\hat{H}_{a(\gamma)_{|v}}f_\gamma = \hat{U}(\varphi_{\gamma,v,v'})\hat{H}_{a(\gamma)_{|v'}}\hat{H}_{a(\gamma)_{|v}}f_\gamma \quad \text{(II.1.3. 16)}$$

for any $v \neq v' \in V(\gamma)$. It follows that

$$\begin{aligned}[\hat{H}(N),\hat{H}(N')]f_\gamma &= \frac{1}{2}\sum_{v\neq v'\in V(\gamma)}[N'(v)N(v')-N(v)N'(v')]\times \\ &\quad \times[\hat{U}(\varphi_{\gamma,v,v'})-\hat{U}(\varphi_{\gamma,v',v})]\hat{H}_{a(\gamma)_{|v'}}\hat{H}_{a(\gamma)_{|v}}f_\gamma \quad \text{(II.1.3. 17)}\end{aligned}$$

where we have used $[\hat{H}_{a(\gamma)_{|v'}},\hat{H}_{a(\gamma)_{|v}}]=0$ for $v \neq v'$ since the derivative operators involved act on disjoint sets of edges.

Expression (II.1.3. 17) is to be compared with the classical formula $\{H(N),H(N'\}=\vec{H}(q^{-1}[(dN)N'-(dN')N])$. The fact that we get a difference between finite diffeomorphism constraint operators looks promising at first because for next neighbour vertices v, v' this could be interpreted as a substitute for the operator $\hat{H}_a$ which somehow had to be written in terms of finite diffeomorphism anyway because we know that the infinitesimal generator dos not exist. Unfortunately there are also contributions from pairs v, v' which are far apart. This we could avoid by specifying the choice function more closely in the sense that the arcs $a_{\gamma,v,e,e'}$ should, for a given vertex v, not depend on all of γ but only on $\gamma_v \subset \gamma$, the subset of γ consisting of all edges incident at v. But still (II.1.3. 17) does not, at least not obviously, resemble the classical calculation too closely because there it is crucial that $\{H(x),H(x')\}\neq 0$ as $x \to x'$ while $[\hat{H}_{a(\gamma)_{|v'}},\hat{H}_{a(\gamma)_{|v}}]=0$ for any $v \neq v'$.

Certainly then for $\Psi \in \mathcal{D}^*_{Diff}, f \in \mathcal{D}$ we have in the URST

$$\Psi([\hat{H}(N),\hat{H}(N')]f) := \lim_{\epsilon\to 0}\lim_{\epsilon'\to 0}\Psi([\hat{H}^\epsilon(N),\hat{H}^{\epsilon'}(N')]f)=0 \quad \text{(II.1.3. 18)}$$

where the limit is again *uniform* in both Ψ, f. But this would not be surprising even if the right hand side would be a manifset quantization of the right and side (with $\vec{H}$ replaced by $\hat{U}(\varphi)-1_{\mathcal{H}^0}$ and ordered to the outmost left). In other words, in the URST we do not see the difference between any operators which are, like (II.1.3. 17), of the form of a difference between two finite diffeomorphism operators ordered to the outmost left times any other operators. Here it proves useful to take the effective operator point of view which in fact *can* detect those differences. Again it requires more work, that is, semiclassical analysis, in order to decide whether the classical limit of the right hand side of (II.1.3. 17) has anything to do with $\vec{H}(q^{-1}([(dN)N'-(dN')N])$. It is worthwhile, however, to point out that (II.1.3. 17) proves the *absence of a strong anomaly*. In other words, if (II.1.3. 18) would not hold, then the quantization that we have proposed would be *mathematically inconsistent*. What is possible though is that (II.1.3. 17) could represent a *weak anomaly* in the sense that the quantum dynamics that $\hat{H}(N)$ generates is *physically inconsistent*, that is, the classical dynamics is not reproduced in the classical limit. This is precisely what has to be analyzed in the future and, if true, to find out how to cure the problem.

To summarize: The constraint algebra of the Hamiltonian constraints among each other is mathematically consistent but possibly has a physical anomaly.

Three remarks are in order:

i)

In [179, 180] the authors prove a statement similar to (II.1.3. 18) on their space $\mathcal{D}^*_\star$. The algebra of their dual constraint operators becomes Abelean for a large class of operators which even classically do not need to be proportional to a diffeomorphism constraint. They then argue that the quantization method proposed here cannot be correct because it either implies a physical anomaly or, even worse, that the (dual of the) quantum metric operator $\hat{q}^{ab}$ vanishes identically.

We disagree with this conclusion for two reasons:

1)

Their limit dual operators are defined by

$$[\hat{H}'(N)\Psi](f) := \lim_{\epsilon\to 0}\Psi(\hat{H}^\epsilon(N)f) \tag{II.1.3. 19}$$

where convergence is only pointwise, that is, for any $\delta > 0$, $\Psi \in \mathcal{D}^*_\star, f \in \mathcal{D}$ there exists $\epsilon(\delta,\Psi,f)$ such that

$$|[\hat{H}'(N)\Psi](f) - \Psi(\hat{H}^\epsilon(N)f)| < \delta \tag{II.1.3. 20}$$

for any $\epsilon < \epsilon(\delta,\Psi,f)$. Thus, while they have blown up $\mathcal{D}^*_{Diff}$ to $\mathcal{D}^*_\star$, their convergence is weaker when restricted to $\mathcal{D}^*_{Diff}$ so that it is not easy to compare the two operator topologies (notice that we can also define a dual operator vial (II.1.3. 19) restricted to $\mathcal{D}^*_{Diff}$ considered as a subspace of $\mathcal{D}^*$, this subspace is just not left invariant so that we cannot compute commutators of duals). However it is clear that the subspace $\mathcal{D}^*_\star$ is a sufficiently small extension of $\mathcal{D}^*_{Diff}$ in order to make sure that a much wider class of operators converges in their topology than the class that we have mind for our topology since our topology roughly requires that $\Psi(\hat{H}^\epsilon(N)f)$ is already independent of ϵ while their topology only requires that the ϵ dependence rests in the smearing functions N which are required to be smooth at vertices.

Therefore our first conclusion is that it is not surprising that in their topology more operators converge.

Next, let us turn to commutators. In our topology, what is required is that the expression $\Psi([\hat{H}^\epsilon(N),\hat{H}^{\epsilon'}(N')]f)$ *just equals zero independently of how large the graph is on which f depends* because we have identified $\hat{H}^\epsilon(N)$ with the continuum operator. In their topology what happens is that unless the operator $\hat{H}(N)$ has also the properties B), C) besides A) then one gets for the commutator an expression of the form (II.1.3. 15) on which one acts with an element $\Psi \in \mathcal{D}^*_\star$ the result of which is that one gets

$$\begin{aligned}([\hat{H}(N')',\hat{H}(N)']\Psi)(f_\gamma) &= \lim_{\epsilon\to 0}\lim_{\epsilon'\to 0}\sum_{v\in V(\gamma)}\sum_{v'\in V(\gamma\cup a^\epsilon(\gamma)_{|v})-V(\gamma)}[N'(v)N(v')-N(v)N'(v')]\times \\ &\quad\times\Psi[\hat{H}_{a^{\epsilon'}(\gamma\cup a^\epsilon(\gamma)_{|v}))_{|v'}},\hat{H}_{a^\epsilon(\gamma)_{|v}}f_\gamma]\end{aligned} \tag{II.1.3. 21}$$

For the same reason as for $\Psi \in \mathcal{D}^*_{Diff}$ each evaluation of Ψ that appears on the right hand side is already independent of ϵ,ϵ' for any $\Psi \in \mathcal{D}^*_\star$ by definition of that space. Therefore the only ϵ,ϵ' dependence rests in the function $[N'(v)N(v')-N(v)N'(v')]$. Now, while each of the roughly $|V(\gamma)|$ $\Psi-$evaluations is nonvanishing, since we take the limit pointwise and the N,N' are smooth, the limit vanishes. If we would not have taken pointwise convergence, then for each finite ϵ,ϵ' we can find f_γ, Ψ such that the right hand side of (II.1.3. 21) takes an arbitrarily large value. The reason for why this happens is that since one of the conditions B), C) does not hold, now the vertices $V(\gamma \cup a^\epsilon(\gamma)_{|v}) - V(\gamma)$ in fact do contribute.

We conclude that their topology is too weak in order to detect even a mathematical anomaly, not to mention a physical anomaly, and suffices even less to select the physically relevant operators. The details of this calculation will appear in [181].

2)

Finally, coming to their second conclusion, we will explicitly display in the next subsection a quantization of $\vec{H}(q^{-1}[(dN)N' - (dN')N])$. Now in their topology, the dual of that operator again annihilates $\mathcal{D}^*_\star$ but this is again only because one takes only pointwise rather than uniform limits. If one tests this operator on a finite graph then, again because there are finitely many contributions each of which evidently proportional to a term of the form $[N'(v)N(v') - N(v)N'(v')]$, the limit must vanish pointwise, however, uniformly it blows up. In particular, this does not show that $\hat{q}^{ab}$ is the zero operator.

ii)

In [182] we find the that claim the action of the Hamiltonian constraint is too local in order to allow for interesting critical points in the renormalization flow of the theory and that therefore the Hamiltonian constraint must be changed drastically if possible at all.

Four comments are appropriate:

First of all the claim is not even technically true, how non-local the operator $\hat{H}(N)$ is depends on our choice function a which builds a new graph around any vertex of a given graph γ and the details of that new graph around v may depend on an *arbitrarily large neighbourhood of* v (where a neighbourhood of degree n can be background independently defined as the set of edges that one can trace within γ if one performs a closed loop with endpoints v using at most n edges). Secondly, as we have said right at the beginning: We are here just exploring the first naive definition of a Hamiltonian constraint, not even the author of [164] believes that the operator proposed gives the final answer. Thirdly, it is unclear what role a renormalization group should play in a diffeomorphism invariant theory, after all renomalization group analysis has much to do with scale transformations (integrating out momentum degrees of freedom above a certain scale) which are difficult to deal with in absence of a background metric. Finally, suppose that we would manage to write down a physically correct Hamiltonian operator of the type of $\hat{H}(N)$. We could order it symmetrically and presumably find a self-adjoint extension. It would then be possible to diagonalize it and in the associated "eigenbasis" the operator would act in an ultralocal way ! Thus any non-local operator can be made ultralocal in an appropriate basis. A good example is given by the Laplace operator in $\mathbb{R}^n$ which is non-local in position space but ultralocal in momentum space. Of course the momentum eigenfunctions are not eigenfunctions but rather distributions and we must take an uncountably infinite linear combination of them (rather, an integral against a sufficiently nice function, that is, a Fourier transform) in order to obtain an L_2 function on which the Laplacian looks rather non-local. Thus, non-locality is hidden in infinite linear combinations which is the reason for why we are working with $\mathcal{D}^*$ rather than with $\mathcal{D}$.

iii)

The proposal for a Hamiltonian constraint whose dual action is restricted to distributions based on Vasiliev invariants [166, 167] also has an Abelean dual algebra. Presumably this also will no longer be the case after strengthening the topology but this must wait until the space of Vasiliev distributions has been turned into a Hilbert space.

II.1.3.4 The Quantum Dirac Algebra

Recall from section II.1.3.3 that in the URST the commutator of two Hamiltonian constraints vanishes : The non-zero operator on $\mathcal{H}^0$ given by $[\hat{H}(N), \hat{H}(N')]$, is indistinguishable from the zero-operator in the URST. We would like to know whether there exists an operator corresponding to $\vec{H}(q^{-1}[(dN)N' - (dN')N])$ and if it is also indistinguishable from the zero operator in the URST. If that would be true, then we could equate the two operators in the URST. Notice that this is still not satisfactory because one cannot test the correctness of the algebraic form of an operator on its kernel, but it is still an important consistency check whether an operator corresponding to $\vec{H}(q^{-1}[(dN)N' - (dN')N])$ exists at all. More explicitly, we wish to study whether we can quantize

$$O(N, N') := \int d^3x (NN'_{,a} - N_{,a}N') q^{ab} V_b \tag{II.1.3. 22}$$

In [119] we answer this question affirmatively, that is, we manage to quantize a regulated operator $\hat{O}^\epsilon(N, N')$ corresponding to (II.1.3. 22) and prove that it converges in the URST to an operator $\hat{O}(N, N')$. We will not derive the operator but merely give its final expression. However, let us point out once more that while H_a and q^{ab} are known not to have well-defined quantizations because the infinitesimal generator of diffeomorphisms does not exist (section I.4.3) and since q^{ab} has the wrong density weight (section II.1.3.1) the combination $\omega_a q^{ab} V_b$ is a scalar density of weight one and therefore has a chance to result in a well-defined operator for any co-vector field ω_a such as $\omega_a = NN'_{,a} - N_{,a}N'$.

Let γ be a graph, $V(\gamma)$ its set of vertices, $v \in V(\gamma)$ a vertex of γ, introduce the triangulation $T(\gamma)$ of section II.1.3 adapted to γ, let Δ be a tetrahedron of that triangulation such that $v(\Delta) = v$, let $\chi_{\epsilon,v}(x)$ be the smoothened out characteristic function of the neighbourhood $U(v)$ (using, for instance, a partition of unity) and finally let $s_I(\Delta)$ be the endpoint of the edge $e_I(\Delta)$ of Δ incident at v. We define a vector field on σ of compact support by

$$\xi^a_{\epsilon,v,\Delta,I}(x) := \chi_{U(v)}(x) \frac{s^a_I(\Delta) - v^a}{\epsilon} \tag{II.1.3. 23}$$

where ϵ^3 is the coordinate volume of $U(v)$ and for any vector field ξ on σ let φ^ξ_t be the one-parameter group of diffeomorphisms that it generates. Let us also introduce the short-hand notation $\hat{V}(v) := \hat{V}(U(v))$. It was shown in [119] that there is a classical object $O_\gamma(N, N')$ which uses the triangulation $T(\gamma)$ and whose limit, as $\gamma \to \sigma$, in the topology of the phase space coincides with (II.1.3. 22). The quantizations of these objects define densely defined operators $\hat{O}(N, N')$ with consistent cylindrical projections $\hat{O}_\gamma(N, N')$ given by their action on functions f_γ cylindrical over a graph γ. The explicit form of these projections is given by

$$\begin{aligned} \hat{O}(N, N') f_\gamma &= -i \frac{16 \epsilon_{ijk} \epsilon_{ilm}}{\hbar \ell_p^2} \sum_{v \in V(\gamma)} \sum_{v(\Delta) = v(\Delta') = v} [\hat{U}(\varphi_\epsilon^{\xi_{\epsilon,v,\Delta',R}} - \mathrm{id}_{\mathcal{H}}] \times \\ &\times \epsilon^{RST} \epsilon^{NPQ} [N(v) N'(s_N(\Delta)) - N(s_N(\Delta)) N'(v)] \times \\ &\times \mathrm{tr}(\tau_j h_{e_P(\Delta)} [h^{-1}_{e_P(\Delta)}, \sqrt{\hat{V}(v)}]) \mathrm{tr}(\tau_k h_{e_Q(\Delta)} [h^{-1}_{e_Q(\Delta)}, \sqrt{\hat{V}(v)}]) \times \\ &\times \mathrm{tr}(\tau_l h_{e_S(\Delta')} [h^{-1}_{e_S(\Delta')}, \sqrt{\hat{V}(v)}]) \mathrm{tr}(\tau_m h_{e_T(\Delta')} [h^{-1}_{e_T(\Delta')}, \sqrt{\hat{V}(v)}]) f_\gamma \end{aligned} \tag{II.1.3. 24}$$

Basically, what happened in the quantization step was that one had to introduce a point splitting which is why one has a double sum over tetrahedra and again factors of $1/\sqrt{\det(q)}$ got absorbed into Poisson brackets which then were replaced by commutators. Notice that in (II.1.3. 24) the square

root of the volume operator appears.
The fact that the combination $[\hat{U}(\varphi_\epsilon^{\xi_{\epsilon,v,\Delta',R}}-\mathrm{id}_{\mathcal{H}}]$ stands *to the left* shows that $\Psi(\hat{O}(M,N')f)=0$ uniformly in $\Psi\in\mathcal{D}^*_{Diff}$ and $f\in\mathcal{D}$ for any N,N'.

II.1.4 The Kernel of the Wheeler-DeWitt Constraint Operator

In [165] it was investigated to what extent one can solve the *Quantum Einstein Equations* for $\Psi\in\mathcal{D}^*_{Diff}$

$$\Psi(\hat{H}(N)f)=0 \qquad \text{(II.1.4. 1)}$$

for all $N\in C^\infty(\sigma), f\in\mathcal{D}$. This section is devoted to an outline of an explicit construction of the complete and rigorous kernel of the proposed operator $\hat{H}(N)$. While $\hat{H}(N)$ will certainly have to be modified in the future, hopefully the methods that we display here will prove useful for other candidates of $\hat{H}(N)$. Notice that these solutions were really the first honest solutions to the Wheeler-DeWitt constraint in full four dimensional quantum general relativity in terms of connections that have appeared in the literature because the result of calculations performed in [98, 100, 99] was of the type zero times infinity. Also, they were the first ones that have non-zero volume and which do not need non-zero cosmological constant.

We first want to give an intuitive picture of the way that the Hamiltonian constraint acts on cylindrical functions. When looking at (II.1.3. 3) and (II.1.3. 4) one realizes the following :
The Euclidean Hamiltonian Constraint operator, when acting on, say, a spin-network state T over a graph γ, looks at each non-planar vertex v of γ and for each such vertex considers each triple of distinct edges $e,e',\tilde{e}$ incident at it. For each such triple, the constraint operator contains three terms labelled by the three possible pairs of edges that one can form from $\{e,e',\tilde{e}\}$. Let us look at one of them, say (neglecting numerical factors)

$$\mathrm{tr}([h_{\alpha(v;e,e')}-h_{\alpha(v;e,e')^{-1}}]h_{\tilde{s}}[h_{\tilde{s}}^{-1},\hat{V}(U(v))])T\ . \qquad \text{(II.1.4. 2)}$$

The notation is as follows : $s,s',\tilde{s}$ are the segments of $e,e',\tilde{e}$ incident at v that end in the endpoints of the three arcs $a(v;,e,e')$ etc., $\alpha(v;e,e')$ is the loop $s\circ a(v;e,e')\circ(s')^{-1}$ and $U(v)$ is any system of mutually disjoint neighbourhoods, one for ech vertex v. For notational simplicity we have dropped the graph label. Let $j,j',\tilde{j}$ be the spins of the edges $e,e',\tilde{e}$ in T. First of all it is easy to see that the piece $h_{\tilde{s}}[h_{\tilde{s}}^{-1},\hat{V}(U_{\epsilon_0}(v))]$ is invariant under a gauge transformation at the endpoint $\tilde{p}$ of $\tilde{s}$. Therefore the state (II.1.4. 2) is also invariant at $\tilde{p}$ and since $\tilde{p}$ is a two-valent vertex this is only possible if the segments $\tilde{s}$ and $\tilde{e}-\tilde{s}$ of $\tilde{e}$ carry the same spin in the decomposition of (II.1.4. 2) into spin-network states T'. But since $\tilde{e}-\tilde{s}$ carries still spin $\tilde{j}$ (no holonomy along $\tilde{e}-\tilde{s}$ appears in (II.1.4. 2)) we conclude that the spin of $\tilde{e}$ is unchanged in T' as compared to T.

However, the same is not true for e,e' : The piece $[h_{\alpha(v;e,e')}-h_{\alpha(v;e,e')^{-1}}]$ is a multiplication operator and raises the spin of $a(v;,e,e')$ from zero to $1/2$ and (II.1.4. 2) decomposes into, in general, four spin-network states T' where the spins of the segments s,s' are raised or lowered in units of $1/2$ as compared to T, that is, they are $j\pm 1/2, j'\pm 1/2$ respectively while the spins of the segments $e-s,e'-s'$ remain unchanged, namely j,j'. All this follows from basic Clebsh-Gordan decomposition theory for $SU(2)$.

Next we look at the remaining piece $\hat{H}(N)+\hat{H}_E(N)$ of the Lorentzian Hamiltonian constraint. Its most important ingredient are the two factors of the operator $\hat{K}$ which, up to a numerical factor, equals $[\hat{V}(\sigma),\hat{H}_E(1)]$. Now as shown in [164], when inserting this operator into (II.1.3. 4) what survives in the term corresponding to the vertex v of the graph is just $[\hat{V}(U(v)),\hat{H}_E(U(v))]$. Thus, since the volume operator does not change any spins, the spin-changing ingredient of the action of the remaining piece of $\hat{H}(N)$ at v are two successive actions of $\hat{H}_E(U(v))$ as just outlined.

In summary, the Hamiltonian constraint operator has an action similar to a fourth order polynomial consisting of creation of annihilation operators. What is being created or annihilated are the spins of edges of a graph (notice that an edge with spin zero is the same as no edge at all).

Let us now look at this action in more detail. We will restrict attention only to the Euclidean piece, for the more complicated full action see [165].
Notice that the Euclidean constraint operator creates edges of a special kind, called *extraordinary edges*, namely the arcs $a = a(v; e, e')$. What is special about them is that they end in planar vertices which are either bi- or tri-valent. If they are tri-valent then, moreover, the vertex is the intersection of the two analytical edges a, e where a just ends on an interiour point of e. Moreover, let e, e' be the edges on which a ends. Then the analytical extensions of e, e' end in at least one point and the two possible earliest of their intersection points away from $a \cap e, a \cap e'$ are, together with these analytical extensions, non-planar vertices of γ. However, not only are these edges special, also the spin they carry is special, namely the arc a carries always spin 1/2. We will continue to call this whole set of extraordinary structures an extraordinary edge.

The special nature of these edges allows to classify the full set of labels $\mathcal{S}$ of spin-network states, called *spin-nets*, as follows. Denote by $\mathcal{S}_0 \subset \mathcal{S}$, called *sources*, the set of spin-nets, corresponding to graphs with no extraordinary edges at all.

From these sources one constructs iteratively derived sets $\mathcal{S}_n(s_0)$, $n = 0, 1, 2, ..$ for each source $s_0 \in \mathcal{N}_0$, called spin-nets of level n based on s_0. Put $\mathcal{S}_0(s_0) := \{s_0\}$ and define $\mathcal{S}_{n+1}(s_0)$ as follows : Take each $s \in \mathcal{S}_n(s_0)$, compute $\hat{H}_E(N)T_s$ for all possible lapse functions N, decompose it into spin-network states and enter the appearing spin-nets into the set $\mathcal{S}_{n+1}(s_0)$.

In [165] it is shown that the sets $\mathcal{S}_n(s_0), \mathcal{S}_{n'}(s_0')$ are disjoint unless $s_0 = s_0'$ and $n = n'$. It is easy to see that the complement of the set of sources $\overline{\mathcal{S}_0} = \mathcal{S} - \mathcal{S}_0$ coincides with the set of derived spin-nets of level greater than zero. Moreover, for each $s \in \mathcal{N}$ there is a unique integer n and a unique source s_0 such that $s \in \mathcal{S}_n(s_0)$.

The purpose for doing all this is, of course, that this classification leads to a simple construction of all rigorous solutions of the Euclidean Hamiltonian constraint based on the observation that

$$\hat{H}_E(N) \cdot \text{span}\{T_s\}_{s \in \mathcal{S}_n(s_0)} \subset \text{span}\{T_s\}_{s \in \mathcal{S}_{n+1}(s_0)} \; . \tag{II.1.4. 3}$$

Since a solution Ψ of (II.1.4. 1) is a diffeomorphism invariant distribution in $\mathcal{D}^*_{Diff}$ we define first $[\mathcal{S}_n(s_0)] := \{[s]\}_{s \in \mathcal{S}_n(s_0)}$ where $[s]$ is the label for the diffeomorphism invariant distribution $T_{[s]}$ (recall section I.4.3). We can now make an ansatz for a basic solution of the form

$$\Psi := \Psi_{[s_0],\vec{n}} := \sum_{k=1}^{N} \sum_{[s] \in [\mathcal{S}_{n_k}(s_0)]} c_{[s]} T_{[s]} \tag{II.1.4. 4}$$

with complex coefficients $c_{[s]}$ which are to be determined from the Quantum Einstein Equations (II.1.4. 1). Now from (II.1.4. 4) it is clear that $\Psi_{[s_0],[\vec{n}]}(\hat{H}_E(N)T_s)$ can be non-vanishing if and only if $[s] \in [\mathcal{S}_{n_k - 1}(s_0)]$ for some $k = 1, .., n$, say $k = l$. Choose a representant $s \in [s]$ and let γ be the graph underlying s and $V(\gamma)$ its set of vertices. We then find, writing $\hat{H}_E(N) = \sum_{v \in V(\gamma)} N(v) \hat{H}_E(v)$, that

$$\Psi_{[s_0],\vec{n}}(\hat{H}_E(N)T_s) = \sum_{[s'] \in [\mathcal{S}_{n_l}(s_0)]} c_{[s']} \sum_{v \in V(\gamma)} N(v) T_{[s']}(\hat{H}_E(v)T_s) \tag{II.1.4. 5}$$

should vanish for any choice of lapse function $N(v)$. Since $N(v)$ can be any smooth function we find the condition that

$$\Psi_{[s_0],\vec{n}}(\hat{H}_E(N)T_s) = \sum_{[s']\in[\mathcal{S}_{n_l}(s_0)]} c_{[s']}T_{[s']}(\hat{H}_E(v)T_s) = 0 \qquad \text{(II.1.4. 6)}$$

should vanish for each choice of the finite number of vertices $v \in V(\gamma)$ and for each of the finite number of spin-nets $s \in \mathcal{S}_{n_l-1}(s_0)$. This follows from the fact that the numbers $T_{[s']}(\hat{H}_E(v)T_s)$ are diffeomorphism invariant and therefore do not actually depend on v itself but only on the diffeomorphism invariant information that is contained in the graph γ together with the vertex v singled out.

Therefore, (II.1.4. 6) is a finite system of linear equations for the coefficients $c_{[s']}$. As the cardinality of the sets $\mathcal{S}_n(s_0)$ exponentially grows with n this system is far from being overdetermined and we arrive at an infinite number of solutions. The most general solution will be a linear combination of the elementary solutions (II.1.4. 6). Qualitatively the same result holds for the Lorentzian constraint [165], however, it is more complicated because coefficients from different levels get coupled and so one gets solutions labelled also by the highest level that was used (possibly one has to allow all levels, that is, the highest level is always infinity). Nevertheless it is remarkable how the solution of the Quantum Einstein Equations is reduced to an exercise in finite-dimensional linear algebra (although the computation of the coefficients $T_{[s']}(\hat{H}_E(v)T_s)$ is far from easy, see, e.g., [183] which, although the authors restrict to tri-valent graphs and $\hat{H}_E(N)$ only, is already rather involved). On the other hand, it is expected that physically interesting solutions will actually be infinite linear combinations of coupled solutions, that is, solutions of infinite level, an intuition coming from [184].

Notice that the solutions (II.1.4. 6) are bona fide elements of $\mathcal{D}^*_{Diff}$ and therefore give, for the first time, rigorously defined solutions to the diffeomorphism and the Hamiltonian constraint of full, four-dimensional Lorentzian Quantum General Relativity in the continuum, subject to the reservation that we still have to prove that the classical limit of this theory in fact is general relativity. One should now organize these solutions into a Hilbert space such that adjointness and canonical commutation relations of full Dirac observables are faithfully implemented. Since group averaging does not work for open algebras, there is no good proposal at this point for how to do that and is a very important open research problem.

II.1.5 Further Related Results

We list here further results that are directly connected to the issues that we have touched upon in this section already.

II.1.5.1 Generator of the Wick Transform

In principle we could dispense with the rigorous construction of the Wick transform since we could work entirely with the operators (II.1.3. 3) and (II.1.3. 4) rather than with the modified ones described in this subsection. However, since the availability of a complex connection representation could be conceptually important in particular when making contact with the path integral formulation, we will make a short digression on available first ansätze for how to get there.

Recall that the generator of the Wick transform $\hat{C}$ is given, up to numerical factors, by $i[\hat{V}(\sigma), \hat{H}_E(1)]$. One would like to invoke the spectral theorem in order to define its exponential and it is therefore motivated to have an at least symmetric operator $\hat{H}_E(1)$. This, however, is not the case the way $\underline{\hat{H}_E(1) \text{ is defined}}$: take for example $s \in \mathcal{S}_0, s' \in \mathcal{S}_1(s)$ such that $< T_{s'}, \hat{H}_E(1)T_s >\neq 0$, but then $< T_s, \hat{H}_E(1)T_{s'} > = 0$ by definition of $\mathcal{S}_0$. Of course, one can symmetrize the operator by defining the matrix elements of the symmetric operator to be the half the sum of the matrix elments of the unsymmetric operator plus the transpose of its complex conjugate. This operator is also well-defined but in [165] we did not succeed in proving existence of self-adjoint extensions of it. What works is the following : one *marks* the extraordinary edges by taking them to be smooth but not analytical. This way it becomes possible to tell whether a given state was obtained by the action of the constraint on a function cylindrical over a piecewise analytical graph. Then the repeated action of the constraint adds always the same smooth extraordinary edge $a(v;e,e')$ to the graph and this turns the operator into a symmetric one when factor-ordering its expression symmetrically. The formalism is not disturbed by the fact that one leaves the purely analytical category of graphs. The so symmetrized operator is free of mathematical anomalies as well but the structure of its solutions becomes more complicated. One can then invoke von Neumann's theorem (that says that if a densely defined symmetric operator commutes on its domain with a conjugation operator that preserves its domain then there exist self-adjoint extensions) to show that $\hat{H}_E(N)$ and in fact also $\hat{C}$ have self-adjoint extensions. (A conjugation operator is a bounded, anti-linear operator which squares to the identity). This method of proof does not work, however, for the constraint $\hat{H}(N)$ because the Lorentzian operator is a sum consisting of two operators which are symmetric and have self-adjoint extensions but it is unclear whether they have extensions to the same domain (the explicit extensions are not even known although it is likely that all operators in question are essentially self-adjoint in which case the answer would be given by their closure).

In any case, we could in principle define self-adjoint operators $\hat{H}_E(N), \hat{C}$ and define the operator $\hat{W}_t := \exp(-t\hat{C})$ by its spectral resolution. Then, according to the philosophy of the Wick transform of section II.1.1 we should analytically continue the operator $\hat{W}_t^{-1}\hat{H}_E(N)\hat{W}_t$ and define it to be the Lorentzian Hamiltonian constraint. The problem is that the spectrum of $\hat{C}$ is far from known and it is not even clear, although extremely likely, that $\hat{H}_E(N)$ and $\hat{C}$ can be extended as self-adjoint operators to the same domain. A method of proof, as sketched in [165], could probably be based on Nelson's analytic vector theorem but the proof was not completed there.

II.1.5.2 Testing the New Regularization Technique by Models of Quantum Gravity

Presently there are two positive tests for the quantization procedure that we applied to the Hamiltonian constraint, namely Euclidean $2+1$ gravity [184] and isotropic and homogeneous BIanchi cosmologies quantized in a non-standard fashion [185].

The first model is a dimensional reduction of $3+1$ gravity which one can formulate also as a quantum theory of $SU(2)$ connections and $su(2)$ electric fluxes with precisely the same algebraic form of all constraints. Hence, one can introduce the full mathematical structure of $\overline{\mathcal{A}}, \mu_0, \mathcal{H}^0$ as well as the quantum constraints $G_j = \mathcal{D}_a E^a_j,\ V_a = F^j_{ab}E^b_j,\ H_E = F^j_{ab}E^a_k E^b_l \epsilon_{jkl}/\sqrt{\det(q)}$ the only difference with the Lorentzian $3+1$ theory being that now indices $a,b,c,..=1,2$ have range in one dimension less and that there is only the Euclidean constraint.

The second model is $3+1$ Lorentzian gravity but instead of performing the usual Killing reduction one looks for (distributional) states in the *full Hilbert space* $\mathcal{H}^0$ of the theory which are compatible with the Killing symmetries of the model.

In both models one then follows step by step the regularization procedure outlined in sections II.1.2,II.1.3. The outcomes are as follows:

- Euclidean 2+1 Gravity
  The quantization of 2+1 general relativity is an exhaustively studied problem (see, e.g., [186, 187, 188, 189, 190, 191, 192], third and fourth references in [103] and references in all of those). Several different quantization techniques have been applied and were shown to give consistent

results. The reader might wonder why 2+1 *Euclidean* quantum gravity should serve as a test model for 3+1 *Lorentzian* quantum gravity. The reason for this is that, as pointed out in [191, 192], the Hamiltonian formulation of 2+1 gravity via connections leads to the non-compact gauge group $SU(1,1)$ for three-metrics of Lorentzian signature while for three-metrics of Euclidean signature we have the same compact gauge group as in Lorentzian 3+1 gravity, namely $SU(2)$. Thus, in order to maximally simulate the 3+1 theory, we should consider Euclidean 2+1 gravity.

However, in order to maximally test the new technique introduced in sections II.1.2, II.1.3 and the constraints of the 3+1 theory one has to develop techniques different from those that people normally employ in 2+1 gravity which make [184] of interest by itself. In particular, it contains a full fledged derivation of the $2+1$ volume operator. The reason is the following:
Pure $2+1$ gravity on a Riemann surface of some fixed genus is a topological field theory, that is, there are only finitely many degrees of freedom. This can be easily seen from the fact that we have six canonical pairs and six first class constraints. When the metric q_{ab} is non-degenerate, the Diffeomorphism and Hamiltonian constraint together are equivalent to the *Flatness Constraint* $C^j := \epsilon^{ab}F^j_{ab} = 0$. Almost exclusively the theory is quantized using C_j rather than V_a, H, see in particular [187] and reference three and four of [103]. But of course we must use V_a, H in order to test the $3+1$ theory appropriately.

The result is that all steps of the quantization programme can be carried up to and including the construction of the full solution space to all constraint. The structure of that solution space is as complicated as in the $3+1$ theory, therefore it is not easy to find a suitable physical inner product. *However, one finds that the full space of solutions to the flatness constraint is contained as a subspace in the space of solutions to the Diffeomorphism and Hamiltonian constraint.* Therefore, the validity of the quantization method is confirmed in this model. Futhermore, the curvature operator F_{ab} (which of course becomes substituted by a holonomy along contractible loop) *must be ordered to the most left.* Thus, ordering is important here, in particular the Hamiltonian constraint is by far not even symmetric in that ordering. An inner product on that subspace of the full solution space is then suggested to be usual product introduced in [187]. The full solution space of the Diffeomorphism and Hamiltonian constraint is much larger which could be related to the fact that it contains a huge number of states with vanishing volume [184], however, this speculation is yet unconfirmed. What to do with zero volume states (degenerate three metrics) in $3+1$ connection quantum gravity has always been a puzzle [186].

- Isotropic, Homogeneous Bianchi Cosmologies
  In an outstanding series of papers [185] Bojowald has introduced a method for embedding the quantum theory of a Killing – or dimensionally reduced model of a given field theory of connections into the quantum theory of the full unreduced theory. Paraphrasing somewhat the procedure, roughly what happens is the following:
  In contrast to the usual mini (midi) superspace quantization procedure of first reducing the classical theory by the Killing symmetry and then quantizing the resulting reduced theory, here one starts with the Hilbert space of the full unreduced theory and imposes the Killing symmetry on states. Since the Killing symmetry group is in a sense a subgroup of the diffeomorhism group it is clear that one gets symmetric (distributional) states by a kind of group averaging procedure together with a natural group averaging inner product. One then "projects" the constraint operators of the full theory, regularized by the same technique as in section II.1.3, on the space of symmetric states. The method is very general but most is known for the isotropic and

homogeneous Bianchi models and has culminated in a series of papers entitled "Loop Quantum Cosmology I – IV" plus successors thereof [185]. The results are indeed spectacular:

1) *Absence of Initial Singularity*
   Let us take as a detector for the big bang singularity in a Friedman – Robertson – Walker model the divergence of the inverse scale factor $1/a$ at $a = 0$. *In quantum theory the big bang singularity evaporates !!!* It is simply not there. The technical reason is as follows: The function $1/a$ is classically a negative power of the volume functional which we can display as the Poisson bracket between a positive power of the volume functional and a holonomy by using the same key identities that were employed in section II.1.2. **Now one simply replaces the Poisson bracket by a commutator exactly as we have done in the full theory and in that very precise sense, the results by Bojowald confirm the validity of our quantization technique of [164]**. The crucial fact is now that the commutator cannot blow up at low volume. Even better, one can show that the quantization of $1/a$ is a positive semi-definite operator which is bounded from above ! Thus, in loop quantum cosmology one of the dreams about quantum gravity, that it cures classical singularities seems to come true.
2) *Rapid Convergence to the Classical Regime*
   One might imagine that a quantization of $1/a$ with such bizarre properties should have a spectrum far off the classcal curve $a \mapsto 1/a$. However, this is far from true: As one can imagine, (linear combinations of symmetric subsets of) spin network states diagonalize the inverse scale factor operator and its spectrum is purely discrete. Already for spins as low as $j > O(10)$ the spectrum lies exaclty on the classical curve.
3) *Discrete Time Evolution*
   The scale factor itself makes a good time observable in the present model. As the solution algorithm for the Hamiltonian constraint of the full theory, section II.1.4, already indicates, the *time evolution becomes discrete* and can be solved in closed form. Amazingly, given initial data one can *quantum evolve through the classical singularity* to negative times.
4) *No Boundary Conditions Necessary*
   When analyzing the solution space of the quantum evolution equation (Wheeler – Dewitt equation) one discovers that there is a *unique* solution with the correct semiclassical properties. Thus, one does not need to carefully adjust the boundary condition on the initial state [186] in order to get a state today with the desired properties, *the theory itself chooses that state*!

One should, of course, consider more examples and transfer these results to the full theory in order to gain confidence into the quantization method, ultimatively a semiclassical analysis is unavoidable, however, these two results described are hopefully promising enough in order to take the proposal for the regularization of the Hamiltonian constraint proposed sufficiently serious.

II.1.5.3 Quantum Poincaré Algebra

In [157] an investigation was started in order to settle the question whether $\mathcal{H}^0$ supports the quantization of the ADM energy surface integral

$$E_{ADM}(N) = -\frac{2}{\kappa}\int_{\partial\sigma} dS_a \frac{N}{\sqrt{\det(q)}} E^a_j \partial_b E^b_j \qquad \text{(II.1.5. 1)}$$

for an asymptotically flat spacetime M (here $\partial\sigma$ corresponds to spatial infinity i^0 in the Penrose diagramme describing the conformal completion of M). It should be stressed that (II.1.5. 1) is the value of the graviatational energy (at unit lapse $N = 1$) only when the constraints are satisfied, otherwise one has to add to (II.1.5. 1) the Hamiltonian constraint $H(N)$. In particular one has to use $H_{ADM}(N) = H(N)+E_{ADM}(N)$ in order to compute the equations of motion. If N is, say of rapid decrease, then $H_{ADM}(N) = H(N)$ generates gauge transformations (time reparameterizations), if it is asymptotically constant then it generates symmetries. There are nine more surface integrals of the type (II.1.5. 1) and together they generate the asymptotic Poincaré algebra. They are the only ten Dirac obeservables known for full, Lorentzian, asymptotically flat gravity in four dimensions. For a discussion of these and related issues, see e.g. [87] and references therein.

In [157] we were only able to cover time translations (II.1.5. 1), spatial translations and spatial rotations. Boosts, which are much harder to define, were not considered there but there is no principal problem to do so. We will focus here only on the quantization of (II.1.5. 1) for reasons of brevity. The method of regularization and quantization completely parallel those displayed in sections II.1.2, II.1.3 and will not be repeated here. The only new element that goes into the classical regularization is the exploitation of the fall off conditions on the classical fields, in particualr that $A = O(1/r^2)$ in an asymptotic radial coordinate. This enables one to replace, effectively, in (II.1.5. 1) $\partial_b E^b_j$ by the gauge invariant quantity $G_j = \mathcal{D}_b E^b_j$ in (II.1.5. 1), that is, the Gauss constraint. At first sight one is tempeted to set it equal to zero. However, a detailed analysis shows that for the Gauss constraint to be functionally differentiable, its Lagrange multiplier must fall off as $1/r^2$ which means that the Gauss constraint does not need to hold at $\partial\sigma$. Thus, it would be physically incorrect to require $G_j = 0$ at $\partial\sigma$, in other words, quantum states do not need to be gauge invariant at $\partial\sigma$ or, put differently, the motions generated by G_j at $\partial\sigma$ are not gauge transformations but symmetries.

The final answer is ($E_{ADM} = E_{ADM}(1)$)

$$\hat{E}_{ADM} f_\gamma = -2m_p \sum_{v\in V(\gamma)\cap\partial\sigma} \frac{\ell_p^3}{\hat{V}_v} R^j_v R^j_v f_\gamma \qquad \text{(II.1.5. 2)}$$

where $R^j_v = \sum_{f(e)=v} R^j_e$, $\hat{V}_v = \lim_{R_v\to\{v\}} \hat{V}(R_v)$ and $x \mapsto R_x$ is an open region valued function with $x \in R_x$. The operator (II.1.5. 2) is defined actually on an extension of $\mathcal{H}^0$ which allows for edges that are not compactly supported. Moreover we must require that 1) for each $v \in \gamma\cap\partial\sigma$ the eigenvalues of $\hat{V}_v$ are non-vanishing and that 2) $e\cap\partial\sigma$ is a discrete set of points for every $e \in E(\gamma)$. We have assumed w.l.g. that all edges with $e\cap\partial\sigma \neq \emptyset$ are of the "up" type with respect to the surface $\partial\sigma$.

Under these assumptions one can show the following:

i) *Positive Semi-Definiteness*
(II.1.5. 2) defines a self-consistent family of essentially self-adjoint, positive semi-definite operators. This is like a quantum positivity of energy theorem but it rests heavily on the two assumptions 1) and 2) made above whose physical origin is poorly understood.

ii) *Fock Space Interpretation*
Since the volume operator is gauge invariant, it follows that it commutes with the Laplacian $\Delta_v = (R^j_v)^2$ and therefore we can simultaneously diagonalize these operators. It is clear that the eigenstates are certain linear combinations of spin-network states and the eigenvalues are of the form $j_v(j_v + 1)/\lambda_v$ (where λ_v is a volume eigenvalue) times m_p. Thus we can complete the intuitive picture that the Hamiltonian constraint gave us: while the constraint changes the spin quantum numbers, the energy is diagonal in it in very much the same way as the annihilation and creation operators of quantum mechanics change the occupation number of

an energy eigenstate. *We may thus interpret the spin quantum numbers as occupation numbers of a non-linear Fock representation.* In quantum field theory we label Fock states by occupation numbers n_k for momentum modes k. Here we have occupation numbers j_e for "edge modes" e.

iii) *Spectral Properties*
The eigenvalues are discrete and unbounded from above but in contrast to the geometry operators there is no energy gap. Rather there is an accumulation point at zero because $[\Delta_v, \hat{V}_v] = 0$ (we can choose the state to be very close to being gauge invariant but to have arbitrarily large volume). This is to be expected on physical grounds because we should be able to detect arbitrarily soft gravitons at spatial infinity.

iv) *Quantum Dirac Observable and Schrödinger equation*
(II.1.5. 2) trivially commutes with all constraints (since diffeomorphisms φ and lapses N that generate gauge transformations are trivial (identity and zero) at $\partial\sigma$) and therefore represents a true quantum Dirac observable. In principle we can now solve "the problem of time" since a physically meaningful time parameter is selected by the one parameter unitary groups generated by $\hat{E}_{ADM}$, in other words, we have a Schrödinger equation

$$-i\hbar\frac{\partial\Psi}{\partial t} = \hat{E}_{ADM}\Psi \tag{II.1.5. 3}$$

Actually in [157] concepts that go beyond $\mathcal{H}^0$ were needed and introduced heuristically. They go under the name "Infinite Tensor Product Extension" and were properly defined only later in [136]. They will be discussed briefly in section II.3.2.

II.2 Extension to Standard Matter

The exposition of section II.1 would be unserious if we would not be able to extend the framework also to matter, at least to the matter of the standard model. This is straightforward for gauge field matter, however for fermionic and Higgs matter one must first develop a background independent mathematical framework [194]. We will discuss the essential steps in the next subsection and then sketch the quantization of the matter parts of the total Hamiltonian constraint in the section after that, see [193] for details.

We should point out that these representations are geared towards a background independent formulation. The matter Hamiltonian operator of the standard model in a background spacetime *is not carried* by these representations. They make sense *only if we couple quantum gravity.* Also, while we did not treat supersymmetric matter explicitly, the following exposition reveals that it is straightforward to extend the formalism to Rarita – Schwinger fields.

II.2.1 Kinematical Hilbert Spaces for Diffeomorphism Invariant Theories of Fermionic and Higgs Fields

First attempts to couple quantum field theories of fermions to quantum general relativity gravity were made in the pioneering work [195]. However, this paper was still written in terms of the complex-valued Ashtekar variables for which the kinematical framework was missing. Later on [196] appeared in which a kinematical Hilbert space for diffeomorphism invariant theories for fermions was proposed which were coupled to arbitrary gauge fields and real-valued Ashtekar variables using the kinematical framework of section I.2. Also, the diffeomorphism constraint was solved there but not the Hamiltonian constraint. However, that fermionic Hilbert space did implement the correct reality conditions for the fermionic degrees of freedom only for a subset of all kinematical observables. In [194] we removed this problem by introducing new fermionic variables, so-called *Grassmann-valued half-densities* and also extended the framework to Higgs fields. This section is accordingly subdivided into one section each for the fermionic and the Higgs sector respectively and in the third section we collect results and define the most general gauge and diffeomorphism invariant states of connections, fermions and Higgs fields.

II.2.1.1 Fermionic Sector

We will take the fermionic fields to be Grassmann-valued, see [197, 198] for a mathematical introduction into these concepts. Furthermore, the Grassmann field $\eta_{A\mu}$ is a scalar with respect to diffeomorphisms of σ which carries two indices, $A, B, C, .. = 1, 2$ and $\mu, \nu, \rho = 1, .., \dim(G)$ corresponding to the fact that it transforms according to the fundamental representation of $SU(2)$ and the defining representation of the compact, connected, unimodular gauge group G of a Yang-Mills gauge theory to which it may couple. This can be generalized to arbitrary representations of $SU(2) \times G$ but we refrain from doing that for the sake of concreteness. Notice that it is no loss of generality to restrict ourselves to only one helicity of the fermion as we can always perform a canonical transformation $(i\bar{\sigma}^{A'}, \sigma_{A'}) \to (i\epsilon^{AB'}\sigma_{B'}, \epsilon_{AB'}\bar{\sigma}^{B'}) =: (i\bar{\eta}^A, \eta_A)$. We will restrict to only one fermionic species in order not to clutter the formulae.

It turns out that the real-valued action in Hamiltonian form for any diffeomorphism invariant theory of fermions is given by

$$S_F = \int_{\mathbb{R}} dt \int_\sigma d^3x (\frac{i}{2}\sqrt{\det(q)}[\bar{\eta}^{A\mu}\dot{\eta}_{A\mu} - \dot{\bar{\eta}}^{A\mu}\dot{\eta}_{A\mu}] - [\text{more}]) \tag{II.2.1. 1}$$

where "more" stands for various constraints and possibly a Hamiltonian and $\det(q)$ is the determinant of the gravitational three-metric which appears because in four spacetime dimensions one needs a metric to define a diffeomorphism invariant theory of fermions. Notice that (II.2.1. 1) is real valued with respect to the usual involution $(\theta_1..\theta_n)^*=\bar\theta_n..\bar\theta_1$ for Grassmann variables $\theta_1,..,\theta_n$ since indices A,μ are raised and lowered with the Kronecker symbol (the involution is just complex conjugation with respect to bosonic variables).

The immediate problem with (II.2.1. 1) is that it is not obvious what the momentum $\pi^{A\mu}$ conjugate to $\eta_{A\mu}$ should be. One strategy would be to integrate the second term in (II.2.1. 1) by parts (the corresponding boundary term being the generator of the associated canonical transformation) and to conclude that it is given by $i\sqrt{\det(q)}\bar\eta^{A\mu}$. However, there is a second term from the integration by parts given by $i\dot{E}^a_i e^i_a\bar\eta^{A\mu}\eta_{A\mu}$ which after a further integration by parts combines with the symplectic potential of the real-valued Ashtekar variables to the effect that A^i_a is replaced by $({}^{\mathbb{C}}A^i_a)=A^i_a-ie^i_a\bar\eta^{A\mu}\eta_{A\mu}$ (recall that E^a_i is the momentum conjugate to A^i_a). This is bad because the connection is now complex-valued and the techniques from section I.2 do not apply any longer so that we are in fact forced to look for another method. The authors of [196] also noticed this subtlety in the following form: If one assumes that the connection is still real-valued while $\pi=i\sqrt{\det(q)}\eta$ is taken as the momentum conjugate to η then one discovers the following contradiction: By assumption we have the classical Poisson bracket $\{\pi(x),A(y)\}=0$. Taking the involution of this equation results in $0=-i\eta(x)\{\sqrt{\det(q)}(x),A(y)\}\neq 0$. If we, however, insert instead of A the complex variable $({}^{\mathbb{C}}A)$ into these equations then in fact there is no contradiction as was shown in [194].

The idea of how preserve the real-valuedness of A^i_a *and* to simplify the reality conditions on the fermions is as follows : Notice that if we define the Grassmann-valued half-density

$$\xi_{A\mu}:=\sqrt[4]{\det(q)}\eta_{A\mu} \tag{II.2.1. 2}$$

then (II.2.1. 1) in fact equals

$$S_F=\int_{\mathbb{R}}dt\int_\sigma d^3x(\frac{i}{2}[\bar\xi^{A\mu}\dot\xi_{A\mu}-\dot{\bar\xi}^{A\mu}\xi_{A\mu}]-[\text{more}]) \tag{II.2.1. 3}$$

without picking up a term proportional to $d\det(q)/dt$. Thus the momentum conjugate to $\xi_{A\mu}$ and the reality conditions respectively are simply given by

$$\pi^{A\mu}=i\bar\xi_{A\mu}\text{ and }(\xi)^*=-i\pi,\ (\pi)^*=-i\xi\ . \tag{II.2.1. 4}$$

The fact that ξ,π are half densities may seem awkward at first sight but it does not cause any immediate problems. Also, recall that "half-density-quantization" is a standard procedure in the theory of geometric quantization of phase spaces with real polarizations [51].

It is in fact possible to base the quantization on the half-density ξ as a quantum configuration variable as far as the solution to the Gauss constraint is concerned. Namely, as has been pointed out by many (see, e.g., [195]) an example for a natural, classical, gauge invariant observable is given by

$$P_e(\xi,A,\underline{A}):=\xi_{A\mu}(e(0))C_1^{A\mu,C\rho}(h_e(A))_{CD}(\underline{\pi}(\underline{h}_e(\underline{A})))_{\rho\sigma}C_2^{D\sigma,B\nu}\xi_{B\nu}(e(1)) \tag{II.2.1. 5}$$

where the notation is as follows : By $(A,h_e,\pi(h_e))$ and $(\underline{A},\underline{h}_e,\underline{\pi}(\underline{h}_e))$ respectively we denote (connection, holonomy along an edge e, irreducible representation evaluated at the holonomy) of the gravitational $SU(2)$ and the Yang-Mills gauge group G respectively. The matrices $C^{A\mu,B\nu}$ are projectors on singlet representations of the decomposition into irreducibles of tensor product representations that appear under gauge transformations on both ends of the path $[0,1]\ni t\to e(t)$

and the irreducible representation $\underline{\pi}$ has to be chosen in such a way that a singlet can occur. For example, if $G = SU(N)$ then we can choose $\underline{\pi}$ to be the complex conjugate of the defining representation. In particular, if $G = SU(2)$ as well we can take $\underline{\pi}$ to be the fundamental representation and $C_1^{A\mu,C\rho} = \epsilon^{AC}\epsilon^{\mu\rho}, C_2^{D\sigma,B\nu} = \delta^{DB}\delta^{\sigma\nu}$. For more general groups we may have to take more than one spinor field at each end of the path in order to satisfy gauge invariance.

All this works fine until it comes to diffeomorphism invariance : notice that the objects (II.2.1. 5) behave strangely under a diffeomorphism φ, namely $\varphi \cdot P_e = P_{\varphi(e)}(J_\varphi(e(0))J_\varphi(e(1)))^{-1/2}$ where $J_\varphi(x) = |\det(\partial\varphi(x)/\partial x)|$ is the Jacobian. Since there are analyticity preserving diffeomorphisms which leave e invariant but such that, say, $J_\varphi(e(0))$ can take any positive value it follows that the average of P_e over diffeomorphisms is meaningless. We are therefore forced to adopt another strategy.

The new idea [194] is to "dedensitize" ξ by means of the δ-distribution $\delta(x,y)$ which itself transforms as a density of weight one in either argument. Let $\theta(x)$ be a smooth Grassmann-valued scalar (we drop the indices $A\mu$) and we *define* $\xi(x)$ not to be a smooth function but rather a distribution (already classically). Let $\delta_{x,y} = 1$ for $x = y$ and zero otherwise (a Kronecker δ, not a distribution). Then on the space of test functions of rapid decrease the distribution $\sqrt{\delta(x,y)\delta(z,y)}$ is well-defined and equals $\delta_{x,z}\delta(x,y)$ [194]. As shown in [194] the following transformations (and corresponding ones for the complex conjugate variables)

$$\theta(x) := \int_\sigma d^3y\sqrt{\delta(x,y)}\xi(y) \tag{II.2.1. 6}$$

$$\xi(x) = \sum_{y\in\sigma} \delta(x,y)\theta(y) \tag{II.2.1. 7}$$

are canonical transformations between the symplectic structures defined by the symplectic potentials $i\int_\sigma d^3x\bar{\xi}(x)\dot{\xi}(x)$ and $i\sum_{x\in\sigma}\bar{\theta}(x)\dot{\theta}(x)$ respectively. Notice that (II.2.1. 6) makes sense precisely when ξ is a distributional half-density and in fact one can show that $\xi = \eta\sqrt[4]{\det(q)}$ will precisely display such a behaviour (at least upon quantization) since $\sqrt{\det(q)}$ becomes an operator valued distribution proportional to the δ distribution (recall the formula for the volume operator). The non-trivial anti-Poisson brackets in either case are given by

$$\{\xi(x),\bar{\xi}(y)\}_+ = -i\delta(x,y) \text{ and } \{\theta(x),\bar{\theta}(y)\}_+ = -i\delta_{x,y} \; . \tag{II.2.1. 8}$$

In summary, we conclude that we can base the quantization of the fermionic degrees of freedom on θ as a configuration variable with conjugate momentum and reality structure given by

$$\pi^{A\mu} = i\bar{\theta}_{A\mu} \text{ and } (\theta)^* = -i\pi, \; (\pi)^* = -i\theta \; . \tag{II.2.1. 9}$$

We now have to develop integration theory. This will be based, of course, on the Berezin "integral" [197, 198]. Let $\mathcal{S}(x)$ be the superspace underlying the $2d$ fermionic configuration degrees of freedom $\theta_{A\mu}(x)$ for any $x \in \Sigma$ where $d = 2\dim(G)$. Of course, all these spaces are just copies of a single space $\mathcal{S}$. This superspace can be turned into a trivial σ-algebra $\mathcal{B}(x)$ consisting of $\mathcal{S}(x)$ and the empty set. On $\mathcal{B}(x)$ one can define a probability "measure" dm_x with the additional property that it is positive on "holomorphic" functions (that is, those which depend $\theta(x)$ only and not on $\bar{\theta}(x)$) in the sense that $\int_{\mathcal{S}} dm_x f(\theta(x))^* f(\theta(x)) \geq 0$ where equality holds if and only if $f = 0$. This measure is given by

$$dm(\bar{\theta},\theta) = \prod_{A\mu}(1 + \bar{\theta}^{A\mu}\theta_{A\mu})d\bar{\theta}^{A\mu}d\theta_{A\mu} \tag{II.2.1. 10}$$

and $dm_x = dm(\bar{\theta}(x),\theta(x))$.
Let now $\overline{\mathcal{S}} := \times_{x\in\sigma}\mathcal{S}_x$ be the fermionic quantum configuration space with σ-algebra $\mathcal{B}$ given by

the *direct product* of the $\mathcal{B}(x)$. The Kolmogorov-theorem [58] for uncountable direct products of probability measures ensures that

$$d\mu_F(\bar\theta,\theta) := \otimes_{x\in\sigma} dm_x \tag{II.2.1. 11}$$

is a rigorously defined probability measure on $\overline{\mathcal{S}}$. It can be recovered as the direct product limit (rather than projective limit) from its finite dimensional joint distributions defined by cylindrical functions. Here a function F on $\overline{\mathcal{S}}$ is said to be cylindrical over a finite number of points $x_1,..,x_n$ if it is a function only of the finite number of degrees of freedom $\theta(x_1),..,\theta(x_n)$ and their complex conjugates, that is, $F(\theta) = f_{x_1,..,x_n}(\bar\theta_1(x_1),\theta_1(x_1),..,\bar\theta_n(x_n),\theta_n(x_n))$ where $f_{x_1,..,x_n}$ is a function on $\mathcal{S}^n$. We then have

$$\int_{\overline{\mathcal{S}}} d\mu_F F = \int_{\mathcal{S}^n} dm(\bar\theta_1,\theta_1)..dm(\bar\theta_n,\theta_n) f_{x_1,..,x_n}(\bar\theta_1,\theta_1,..,\bar\theta_n,\theta_n)\ . \tag{II.2.1. 12}$$

Basic cylindrical functions are the fermionic vertex functions. These are defined as follows : order the labels $A\mu$ from 1 to $2d$ and denote them by $i,j,k,..$ (confusion with the $su(2)$ labels should not arise). Denote by I an array $1 \le i_1 < .. < i_k \le 2d$ and define $|I| = k$ in this case (confusion with the Lie(G) or spin-network labels should not arise). Then for each set of distinct points $v_1,..,v_n$ we define

$$F_{\vec v,\vec I} = \prod_{l=1}^{n} F_{v_l,I_{v_l}},\ \ F_{v_l,I_{v_l}} = \prod_{j=1}^{|I_{v_l}|} \theta_{i_j(v_l)}(v_l) \tag{II.2.1. 13}$$

Is this the correct measure, that is, are the adjointness relations $\hat\pi^\dagger = -i\hat\theta,\ \hat\theta^\dagger = -i\hat\pi$ and the canonical anti-commutation relations $[\hat\theta_{A\mu}(x),\hat\pi^{B\nu}(y)]_+ = i\hbar\delta_A^B\delta_\mu^\nu\delta_{x,y}$ faithfully implemented ? It is sufficient to check this to be the case on cylindrical subspaces if we represent $\hat\theta(x)$ as a multiplication operator and $\hat\pi(x)$ as $i\hbar\partial^l/\partial\theta(x)$ where the superscript stands for the left *ordinary* derivative (not a functional derivative). In fact, the measure $d\mu_F$ is uniquely selected by these relations given the representation just as in the case of the theory of distributional connections $\overline{\mathcal{A}}$. Also, it is trivially diffeomorphism invariant since the integrals of a function cylindrical over n points and of its diffeomorphic image coincide.

In summary, the correct kinematical Fermion Hilbert space is therefore defined to be $\mathcal{H}_F := L_2(\overline{\mathcal{S}}, d\mu_F)$. It follows immediately from these considerations that the quantum fermion field *at a point* (i.e. totally unsmeared) becomes a densely defined operator. This seems astonishing at first sight but it is only a little bit more surprising than to assume that Wilson loop operators, the quantum connection being smeared in one direction only, are densely defined. When quantizing diffeomorphism invariant theories which lack a background structure one has to give up standard representations and construct new ones.

II.2.1.2 Higgs Sector

It turns out that it is also not possible to combine the well-developed theory of Gaussian measures for scalar field theories with diffeomorphism invariance in order to obtain a kinematical framework for diffeomorphism invariant theories of Higgs fields. The basic obstacle is that a Gaussian measure is completely defined by its covariance which, however, depends on a background structure (see [193] for a detailed discussion of this point). We are therefore again led to a new non-standard representation.

In the following we restrict ourselves to real-valued Higgs-fields ϕ_I which transform according to the adjoint representation of G. Other cases can be treated by similar methods. We also allow for scalar fields (without internal degrees of freedom).
Since in the previous subsection we already got used to dealing with representations for which the

quantum configuration field at a point becomes a well-defined quantum operator it is perhaps not so awkward anymore to do the same for the Higgs field. Actually, we are not going to deal with ϕ_I itself but with the *point-holonomies*, which also play a crucial role in Bojowald's series [185]

$$U_x(\phi) := \exp(\phi_I(x)\tau_I) \tag{II.2.1. 14}$$

where τ_I denotes a basis of the Lie algebra Lie(G) of the Yang-Mills gauge group. The name stems from the fact that under a gauge transformation $g(x)$ at x we have that $U(x) \to \mathrm{Ad}_{g(x)}(U(x))$ which is precisely the transformation behaviour of a holonomy $\underline{h}_e$ starting at x in the limit of vanishing edge length. In the case of a simple scalar field we define $U_x = e^{i\phi(x)}$. These variables play a role similar to the Wilson loop variables in lattice gauge theory [183] and it is understood that any action written in terms of ϕ_I should be rewritten in terms of the $U(x)$ in analogy to the replacement of the Yang-Mills action by the Wilson action.

This analogy with holonomies suggests a step by step repetition of the Ashtekar-Isham-Lewandowski framework of section I.2 [194] which we are going to sketch below. Before doing that we must decide on the elementary variables. With ϕ_I being a scalar its conjugate momentum p^I is a scalar density of weight one. Therefore the integrated quantity

$$p^I(B) := \int_B d^3x p^I(x) \tag{II.2.1. 15}$$

for any open region B in σ is diffeomorphism covariantly defined and the formal Poisson brackets $\{p^I(x), \phi_J(y)\} = \delta^I_J \delta(x,y)$ translate into

$$\{p^I(B), U_x\} = \chi_B(x)\frac{1}{2}[\tau_I U_x + U_x \tau_I] \tag{II.2.1. 16}$$

(in order to see this one must regularize U_x as in [194] and then remove the regulator). The other elementary Poisson bracket is $\{U_x, U_y\} = 0$. Actually one has to generalize the Poisson algebra to the Lie algebra of functions on smooth ϕ_I's and vector fields thereon just as in the case of connections in order to obtain a true Lie algebra which one can quantize. Finally, $p^I(B)$ is real-valued and U_x is G-valued.

The construction of a quantum configuration space $\overline{\mathcal{U}}$ and a diffeomorphism invariant measure $d\mu_U$ thereon now proceeds just in analogy with section I.2 :
A Higgs vertex function $H_{\vec{v},\vec{\pi},\vec{\mu},\vec{\nu}}$ is just given by

$$H_{\vec{v},\vec{\pi},\vec{\mu},\vec{\nu}} = \prod_{k=1}^{n} \sqrt{d_{\pi_k}}(\pi_k(U(v_k)))_{\mu_k\nu_k} \tag{II.2.1. 17}$$

where π_k are chosen from a complete set of irreducible, inequivalent representations of G and $v_1, .., v_k$ are distinct points of σ.
Consider the Abelian C* algebra given by finite linear combinations of Higgs vertex functions and completed in the sup-norm over the set of smooth Higgs fields $\mathcal{U}$. Then $\overline{\mathcal{U}}$, the quantum configuration space of distributional Higgs fields, is the spectrum of that algebra equipped with the weak* topology (Gel'fand topology).

The characterization of the spectrum is as follows : points $\bar{\phi}$ in $\overline{\mathcal{U}}$ are in one to one correspondence with the set Fun(σ, G) of g-valued functions on σ, the correspondence being given by $\bar{\phi} \leftrightarrow U_{\bar{\phi}}$ where $\sqrt{d_{\pi_0}}(U_{\bar{\phi}})_{\mu\nu}(v) = \bar{\phi}(H_{v,\pi_0,\mu,\nu})$ and π_0 is the fundamental representation of G.

Again, since the spectrum is a compact Hausdorff space one can define a regular Borel probability measure μ on it through positive, normalized, linear functionals Γ on the set of continuous functions f thereon, the correspondence being given by $\Gamma(f) = \int_{\overline{\mathcal{U}}} d\mu f$. We define the measure μ_U by

$$\Gamma_{\mu_U}(H_{\vec{v},\vec{\pi},\vec{\mu},\vec{\nu}}) = \begin{cases} 1 & H_{\vec{v},\vec{\pi},\vec{\mu},\vec{\nu}} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{II.2.1. 18}$$

and one easily sees that this measure is just the Haar measure on G^n for functions cylindrical over n distinct points. In particular, the Higgs vertex functions form a complete orthonormal basis by an appeal to the Peter&Weyl theorem. The measure μ_U can be shown [194] to be concentrated on nowhere continuous Higgs fields, in particular $\mu_U(\mathcal{U}) = 0$.

Finally, $\hat{U}(x)$ is just a multiplication operator on cylindrical functions and if we replace p^I by $-i\hbar\delta/\delta\phi_I$ then we find for a function $F = f_{\vec{v}}$ cylindrical over n points $\vec{v}$ that $\hat{p}^I(B)F = -i\hbar\sum_{k=1}^n \chi_B(v_k) X_v^I f_{\vec{v}}$ where $X_v^I = X^I(U(v))$, $X^I(g) = \frac{1}{2}[X_R^I(g) + X_L^I(g)]$ and X_L, X_R are, respectively, left and right invariant vector fields on G. This representation shows that the canonical commutation relations as well as the adjointness relations are faithfully implemented and that the appropriate kinematical Higgs field Hilbert space can be chosen to be $\mathcal{H}_U := L_2(\overline{\mathcal{U}}, d\mu_U)$.

II.2.1.3 Gauge and Diffeomorphism Invariant Subspace

We now put everything together to arrive at the complete solution to the Gauss and Diffeomorphism constraint for quantum gravity coupled to gauge fields, Higgs fields and Fermions.

We begin with the kinematical Hilbert space

$$\mathcal{H} = L_2(\overline{\mathcal{A}}_{SU(2)}, d\mu_0^{SU(2)}) \otimes L_2(\overline{\mathcal{A}}_G, d\mu_0^G) \otimes L_2(\overline{\mathcal{F}}, d\mu_F) \otimes L_2(\overline{\mathcal{U}}, d\mu_U) \tag{II.2.1. 19}$$

and now consider its subspace consisting of gauge invariant functions. A basis of such functions is labelled by a graph γ, a labelling of its edges e by spins j_e and colours c_e corresponding to irreducible representations of $SU(2)$ and G and a labelling of its vertices v by an array I_v, another colour C_v and two projectors p_v, q_v. The array I_v indicates a fermionic dependence at v by F_{v,I_v} and C_v stands for an irreducible representation of G evaluated at $U(v)$. Finally, decompose the tensor product of irreducible representations of $SU(2)$ given by the fundamental representations corresponding to F_{v,I_v} and the representations π_{j_e} for those edges e incident at v and project with p_v on a singlet that appears. Likewise, decompose the tensor product of irreducible representations of G given by the fundamental representations corresponding to F_{v,I_v}, the representations $\underline{\pi}_{c_e}$ for those edges e incident at v and the representation $\underline{\pi}_{C_v}$ and project with q_v on a singlet that appears.

The result is a gauge invariant state $T_{\gamma,[\vec{j},\vec{I},\vec{p}],[\vec{c},\vec{C},\vec{q}]}$ called a *spin-colour-network state* extending the definition of a purely gravitational spin-network state. Consider the action $\overline{\mathcal{G}}$ of the gauge group $SU(2)\times G$ on all distributional fields. Then the spin-colour network states contain the space of gauge invariant functions which is the same as the Hilbert space

$$\mathcal{H} = L_2([\overline{\mathcal{A}}_{SU(2)} \times \overline{\mathcal{A}}_G \times \overline{\mathcal{F}} \times \overline{\mathcal{U}}]/\overline{\mathcal{G}}, d\mu_0^{SU(2)} \otimes d\mu_0^G \otimes d\mu_F \otimes d\mu_U), \tag{II.2.1. 20}$$

that is, the L_2 space on the moduli space.

To get the solution to the diffeomorphism constraint one considers the spaces $\mathcal{D}_{SU(2)}, \mathcal{D}_G, \mathcal{D}_F, \mathcal{D}_U$ of smooth cylindrical functions (smooth in the sense of the nuclear topology of $SU(2)^n, G^n, \mathcal{S}^n, G^n$ respectively) and their corresponding algebraic duals. Then we form the gauge invariant subspaces of the spaces

$$\mathcal{D} := \mathcal{D}_{SU(2)} \times \mathcal{D}_G \times \mathcal{D}_F \times \mathcal{D}_U \text{ and } \mathcal{D}^* := \mathcal{D}^*_{SU(2)} \times \mathcal{D}^*_G \times \mathcal{D}^*_F \times \mathcal{D}^*_U\,. \tag{II.2.1. 21}$$

Now the spin-colour-network states span the invariant subspace of $\mathcal{D}$ and the diffeomorphism group acts unitarily by

$$\hat{U}(\varphi)T_{\gamma,[\vec{j},\vec{I},\vec{p}],[\vec{c},\vec{C},\vec{q}]} = T_{\varphi(\gamma),[\vec{j},\vec{I},\vec{p}],[\vec{c},\vec{C},\vec{q}]} \tag{II.2.1. 22}$$

and similar as in the purely gravitational case we get diffeomorphism invariant distributions in $\mathcal{D}^*$ by group judiciously averaging the action (II.2.1. 22).

II.2.2 Quantization of Matter Hamiltonian Constraints

We will restrict ourselves, for the sake of clarity, to only one kind of matter, namely pure Yang-Mills fields for a compact, connected gauge group G. See [193] for the general case.

To anticipate the result, what we find is that certain ultraviolet divergences, which appear when we consider Yang-Mills fields propagating on a background spacetime, *disappear* when we let the spacetime metric fluctuate as well. We do not claim that this proves finiteness of quantum gravity because, first, we must prove that the quantum theory constructed has general relativity as its classical limit, secondly, besides the Hamiltonian constraint we also must show that quantizations of classical observables of the theory are finite and, thirdly, we must establish that those operators remain non-singular upon passing to the physical Hilbert space. However, these are first promising indications anyway.

We will first explain how this works in canonical quantum gravity for the Yang-Mills field and then describe the general mechanism.
The canonical pair coordinatizing the Yang-Mills phase space is given by $(\underline{E}^a_I, \underline{A}^I_a)$ with symplectic structure formally given by

$$\{\underline{E}^a_I(x), \underline{A}^J_b(y)\} = \delta^a_b \delta^J_I \delta(x,y) \tag{II.2.2. 1}$$

where as before $I, J, K, ... = 1,.., \dim(G)$ denote $L(G)$ indices. The contribution of the Yang-Mills field to the Hamiltonian constraint turns out to be

$$H_{YM} = \frac{q_{ab}}{2Q^2\sqrt{\det(q)}}[\underline{E}^a_I \underline{E}^b_I + \underline{B}^a_I \underline{B}^b_I] \tag{II.2.2. 2}$$

where Q is the Yang-Mills coupling constant, $\underline{B}^a_I := \frac{1}{2}\epsilon^{abc}\underline{F}^I_{bc}$ the magnetic field of the connection $\underline{A}^I_a$ and $\underline{F}^I_{ab}$ its curvature. The integrated form is given by $H_{YM}(N) = \int_\sigma d^3x N H_{YM}$ where N is the lapse function.

In a background spacetime, say in Minkowski space ($N=1, q_{ab} = \delta_{ab}$) the integrated Hamiltonian constraint becomes just the Hamiltonian of the theory. Let us see what happens if we try to quantize this field theory propagating in Minkowski space non-perturbatively. It will be enough to consider Maxwell-Theory. In order not to have to worry about infrared divergencies, for the sake of the argument, we will add for this paragraph a mass term $m^2 q^{ab}\sqrt{\det(q)}\underline{A}_a\underline{A}_b$ to the Hamiltonian density so that we are actually looking at the Proca field. We can then define a Hilbert space $L_2(\mathcal{S}', d\mu_G)$ where μ_G is some Gaussian measure on the space of tempered distributional connections. The reality conditions and the canonical commutation relations are satisfied if we let $\hat{\underline{A}}_a$ act by multiplication and $\hat{\underline{E}}^a = -i\hbar Q^2 \delta/\delta\underline{A}_a + F^a$ where F^a is a function of the connection chosen in such a way that $\hat{\underline{E}}^a$ is formally a self-adjoint operator (for instance if μ_G is the white noise measure then $F^a = \lambda\underline{A}_a$ for some constant λ).

Let us try to compute the ground state of the Hamiltonian. Since its density is proportional to

$$-[\frac{\delta}{\delta\underline{A}_a} + \lambda F^a]^2 + \underline{A}_a D \underline{A}_a$$

for a positive, invertible differential operator D one will make the ansatz

$$\Psi = \exp(-\frac{1}{2}\int d^3x[\lambda(\underline{A}_a)^2 + \underline{A}_a\sqrt{D}\underline{A}_a])$$

but

$$\int d^3x\hat{H}(x)\Psi \propto \Psi\cdot(\int d^3x)\cdot[\sqrt{D}_x\delta(x)]_{x=0}$$

diverges. This divergence is still rather harmless since it can be removed by factor ordering but in the interacting case things get worse.
What is the origin of the divergence ? As we have already pointed out earlier in this review, it is rooted in the fact that (II.2.2. 2) becomes a density of weight *two* if we take q_{ab} to be non-dynamical, for instance a constant as in the case of Minkowski space. This is because $\underline{E}^a$ has density weight one whether or not we couple gravity. The density weight shows up, for example, in the evaluation of the two functional derivatives at one and the same point, giving a meaningless result since the Lebesgue measure d^3x can only absorb one of the resulting δ distributions. On the other hand, if we treat q_{ab} as a dynamical field then (II.2.2. 2) adopts a density weight of *one* again because $\sqrt{\det(q)}$ has density weight one and is in the denominator. As we have seen in section II.1.3, the fact that $\sqrt{\det(q)}$ is in the denominator does not need to prevent us from being able to give rigorous meaning to an operator corresponding to $H_{YM}(N)$. We have already seen that gravity regulates itself in this restricted sense.

In the next subsection we explain the quantization of $H_{YM}(N)$ and in the subsequent one we will explain the general scheme how coupling gravity is able to regulate certain ultra-violet divergences.

II.2.2.1 Quantization of Einstein-Yang-Mills-Theory

We will focus first on the electric part of $H_{YM}(N)$ which we write in the form

$$H_{YM,el}(N) = \frac{1}{2Q^2}\int d^3xN\frac{[e_a^i\underline{E}_I^a]}{\sqrt{\det(q)}}[e_b^i\underline{E}_I^b] \tag{II.2.2. 3}$$

where Q is the Yang-Mills coupling constant. Using the same notation as in section II.1 we can also write this as

$$H_{YM,el}(N) = \frac{1}{8\kappa^2Q^2}\int d^3xN(x)\frac{[\{A_a^i(x),V(R_x)\}\underline{E}_I^a(x)]}{\sqrt{\det(q)}(x)}[\{A_b^i(x),V(R_x)\}\underline{E}_I^b(x)] \tag{II.2.2. 4}$$

Since $\underline{E}_I^a = \frac{1}{2}\epsilon^{abc}\underline{e}_{bc}^I$ is Hodge dual to a two-form $\underline{e}^I$ we can also write this as

$$H_{YM,el}(N) = \frac{1}{8\kappa^2Q^2}\int d^3xN(x)\frac{[\{A_a^i(x),V(R_x)\}\underline{E}_I^a(x)]}{\sqrt{\det(q)}(x)}[\{A_i(x),V(R_x)\}\wedge\underline{e}_I(x)] \tag{II.2.2. 5}$$

which suggests to approximate the integral by a Riemann sum utilizing a triangulation of σ as in section II.1.2. Using the same notation as there we get

$$\begin{aligned} H^\epsilon_{YM,el}(N) &= \frac{1}{8\kappa^2Q^2}\sum_{\Delta\in T(\epsilon)}N(v(\Delta))\frac{[\{A_a^i(v(\Delta)),V(R_{v(\Delta)})\}\underline{E}_I^a(v(\Delta))]}{\sqrt{\det(q)}(v(\Delta))}\times \\ &\times\ \epsilon^{LMN}[\mathrm{tr}(\tau_i h_{e_L(\Delta)}\{h^{-1}_{e_L(\Delta)},V(R_{v(\Delta)})\})E_I(S_{MN}(\Delta))] \end{aligned} \tag{II.2.2. 6}$$

where we have used that $S_{MN}(\Delta)$ is any oriented triangular surface with boundary $e_M(\Delta)\circ a_{MN}(\Delta)\circ e_N(\Delta)^{-1}$.
We now apply the same trick that we used already in previous sections: Let $\chi_{\epsilon,x}(y)$ be the characteristic function of a box $U_\epsilon(x)$ with coordinate volume ϵ^3 and centre x. Then

$$\begin{aligned} V(U_\epsilon(x)) &= \epsilon^3\sqrt{\det(q)}(x)+o(\epsilon^4) \text{ and} && \text{(II.2.2. 7)}\\ \int \chi_{\epsilon,x}(y)\frac{[\{A_i(y),V(R_y)\}\wedge \underline{e}_I(y)]}{\sqrt{V(U_\epsilon(y))}} &= \epsilon^3\frac{[\{A_a^i(x),V(R_x)\}\underline{E}_I^a(x)]}{\sqrt{V(U_\epsilon(x))}}+o(\epsilon^3) && \text{(II.2.2. 8)}\end{aligned}$$

which allows us to replace (II.2.2. 6) by

$$\begin{aligned} H^\epsilon_{YM,el}(N) &= \frac{1}{2\kappa^2Q^2}\sum_{\Delta,\Delta'\in T(\epsilon)} N(v(\Delta))\chi_{\epsilon,v(\Delta)}(v(\Delta'))\epsilon^{LMN}\epsilon^{RST}\times\\ &\times \frac{\mathrm{tr}(\tau_i h_{e_L(\Delta)}\{h^{-1}_{e_L(\Delta)},V(R_{v(\Delta)})\})E_I(S_{MN}(\Delta))}{2\sqrt{V(U_\epsilon(v(\Delta)))}}\times\\ &\times \frac{\mathrm{tr}(\tau_i h_{e_R(\Delta')}\{h^{-1}_{e_R(\Delta')},V(R_{v(\Delta')})\})E_I(S_{ST}(\Delta'))}{2\sqrt{V(U_\epsilon(v(\Delta')))}} \qquad \text{(II.2.2. 9)}\end{aligned}$$

Again, the region-valued function $x\to R_x$ is completely arbitrary up to this point and if we choose $R_x=U_\epsilon(x)$ then we obtain the final formula

$$\begin{aligned} H^\epsilon_{YM,el}(N) &= \frac{1}{2\kappa^2Q^2}\sum_{\Delta,\Delta'\in T(\epsilon)} N(v(\Delta))\chi_{\epsilon,v(\Delta)}(v(\Delta'))\epsilon^{LMN}\epsilon^{RST}\times\\ &\times [\mathrm{tr}(\tau_i h_{e_L(\Delta)}\{h^{-1}_{e_L(\Delta)},\sqrt{V(U_\epsilon(v(\Delta)))}\})E_I(S_{MN}(\Delta))]\times\\ &\times [\mathrm{tr}(\tau_i h_{e_R(\Delta')}\{h^{-1}_{e_R(\Delta')},\sqrt{V(U_\epsilon(v(\Delta')))}\})E_I(S_{ST}(\Delta'))] \qquad \text{(II.2.2. 10)}\end{aligned}$$

in which the $1/\sqrt{\det(q)}$ was removed from the denominator and so qualifies as the starting point for the quantization. The pointwise limit of (II.2.2. 10) on the phase space gives back (II.2.2. 2) for *any* triangulation.

The theme repeats : in order to arrive at a well-defined result on a dense set of vectors given by functions cylindrical over graphs γ one must adapt the triangulation to the γ in question. The limit of (II.2.2. 10) with respect to the so obtained $T(\epsilon,\gamma)$ gives still back (II.2.2. 2). The only new ingredient of the triangulation as compared to the one outlined in section II.1.3 is that, at fixed ϵ, we deform the surfaces $S_{MN}(\Delta)$, controlled by a further parameter δ, to the effect that $\lim_{\delta\to 0} S_{MN}(\Delta,\delta)=S_{MN}(\Delta)$ and at finite δ the edge $e_L(\Delta)$, $\epsilon^{LMN}=1$ is the only one that intersects $S_{MN}(\Delta,\delta)$ transversally. This can be achieved by detaching $S_{MN}(\Delta)$ slightly from $v(\Delta)$ and otherwise choosing the shape of $S_{MN}(\Delta)$ appropriately. After replacing Poisson brackets by commutators times $1/(i\hbar)$ and the Yang-Mills electric field by $-i\hbar Q^2$ times functional derivatives we first get a family of operators $\hat{H}^{\epsilon,\delta}_{YM,el}(N)_\gamma$, the limit $\delta\to 0$ of which, in the topology of smooth connections, converges to a family of operators $\hat{H}^\epsilon_{YM,el}(N)_\gamma$ which can be extended to all of $\overline{\mathcal{A}}$. One verifies that this family of operators, for sufficiently small ϵ depending on γ qualifies as the set of cylindrical projections of an operator $\hat{H}^\epsilon_{YM,el}(N)$ and the limit $\hat{H}_{YM,el}(N)$ as $\epsilon\to 0$ in the URST exists and is given by $\hat{H}^{\epsilon_0}_{YM,el}(N)$ for any arbitrary but fixed $\epsilon_0>0$. We give the final result

$$\hat{H}_{YM,el}(N)f_\gamma = -\frac{m_p\alpha_Q}{2\ell_p^3}\sum_{v\in V(\gamma)}\sum_{v(\Delta)=v(\Delta')=v}\frac{N(v)}{E(v)^2}\mathrm{tr}(\tau_i h_{e_M(\Delta)}[h^{-1}_{e_M(\Delta)},\sqrt{\hat{V}(U_{\epsilon_0}(v(\Delta)))}])\times$$

$$\times \quad \mathrm{tr}(\tau_i h_{e_N(\Delta')}[h^{-1}_{e_M(\Delta')}, \sqrt{\hat{V}(U_{\epsilon_0}(v(\Delta')))}])\underline{R}^I_{e_M(\Delta)}\underline{R}^I_{e_N(\Delta')}\, f_\gamma \qquad \text{(II.2.2. 11)}$$

where the Planck mass $m_p = \sqrt{\hbar/\kappa}$ and the dimensionless fine structure constant $\alpha_Q = \hbar Q^2$ have peeled out (in our notation, Q^2 has the dimension of $1/\hbar$) while the Planck volume ℓ_p^3 in the denominator makes the rest of expression dimensionless. As before, $\underline{R}^I_e = \underline{R}^I(\underline{h}_e)$ and $\underline{R}^I(\underline{g})$ is the right invariant vector field on G and $\underline{h}_e$ is the holonomy of $\underline{A}$ along e. Expression (II.2.2. 11) is manifestly gauge invariant and diffeomorphism covariant.

Notice that, expectedly, (II.2.2. 11) resembles (minus) a Laplacian. Indeed, one can show [193] that $\hat{H}_{YM,el}(N=1)$ is an essentially self-adjoint, positive semi-definite operator on $\mathcal{H}$. In particular, (II.2.2. 11) is densely defined and does not suffer from any singularities, it is *finite* ! This extends to the magnetic part of the Yang-Mills Hamiltonian whose action on cylindrical functions is given by

$$\begin{aligned}
\hat{H}_{YM,mag}(N)f_\gamma &= -\frac{m_p}{2\alpha_Q(12N^2\ell_p^3}\sum_{v\in V(\gamma)}\sum_{v(\Delta)=v(\Delta')=v}\frac{N(v)}{E(v)^2}\epsilon^{LMN}\epsilon^{RST}\times \\
&\times \quad \mathrm{tr}(\tau_i h_{e_L(\Delta)}[h^{-1}_{e_L(\Delta)}, \sqrt{\hat{V}(U_\epsilon(v(\Delta)))}])\times \qquad \text{(II.2.2. 12)} \\
&\times \quad \mathrm{tr}(\tau_i h_{e_R(\Delta')}[h^{-1}_{e_R(\Delta')}, \sqrt{\hat{V}(U_\epsilon(v(\Delta')))}])\mathrm{tr}(\underline{\tau}_I\underline{h}_{\alpha_{MN}(\Delta)})\mathrm{tr}(\underline{\tau}_I\underline{h}_{\alpha_{ST}(\Delta')})\, f_\gamma
\end{aligned}$$

(we use the normalization $\mathrm{tr}(\underline{\tau}_I\underline{\tau}_J) = -\delta_{IJ}/N$ for the normalization of the generators of Lie(G)). Notice the non-perturbative dependence of (II.2.2. 12) on the fine structure constant.

In summary, the Yang-Mills contribution to the Hamiltonian constraint can be densely defined on $\mathcal{H}$. We can see explicitly the regularizing role that the gravitational quantum field has played in the quantization process : the volume operator acts only at vertices of a graph and therefore also restricts the Yang-Mills Hamiltonian to an action at those points. Therefore, *the volume operator acts as an Infra-Red-Cutoff* ! Next, the divergent factor $1/\epsilon^3$ stemming from the point-splitting of the two Yang-Mills electric fields was *absorbed* by the volume operator which *must happen* in order to preserve diffeomorphism covariance as the point splitting volume should not be measured by the coordinate background metric but by the dynamical metric itself. Therefore, *the volume operator also acts as an Ultra-Violet-Cutoff* ! The volume operator thus plays a key role in the quantization process which is why a more detailed knowledge about its spectrum would be highly desirable.

One can verify that the Quantum Dirac algebra of the complete Hamiltonian constraint $\hat{H}(N) = \hat{H}_{Einstein}(N) + \hat{H}_{YM}(N)$ closes in a similar fashion as outlined in section II.1.4. As shown in [193], this extends to the Fermionic and Higgs sector as well.

That all of this is not coincidence will be the subject of the next subsection.

II.2.2.2 A General Quantization Scheme

Looking at what happened in sections II.1.3 and II.2.2.1 it seems that one can quantize any Hamiltonian constraint which is a scalar density of weight one in such a way that it is densely defined. Indeed, in [193] we gave a proof for this which we sketch below (we restrict ourselves here to non-fermionic matter and to $D=3$ spatial dimensions for the sake of clarity). It applies to any field theory in any dimension $D \geq 2$ which is given in Hamiltonian form, that is, any generally covariant field theory deriving from a Lagrangian (for theories including higher derivatives as in higher derivative gravity [199] or as predicted by the effective action of string theory [13] one can apply the Ostrogradsky method [200] to bring it into Hamiltonian form).

Suppose then that we are given a scalar density $H(x)$ of weight one. Without loss of generality we can assume that all the momenta P of the theory are tensor densities of weight one and act by functional derivation with respect to the configuration variables Q which are associated dual tensor densities of weight zero. By contracting them with triad and co-triad fields we obtain new canonical variables without tensor indices but with $su(2)$ indices. The corresponding canonical transformation is generated by a functional which changes the definition of the real-valued connection variable A_a^i but preserves its real-valuedness and thus does not spoil the kinematical Hilbert space of section I.2. Spatial covariant derivatives are then with respect to A_a^i.

The general form of this density $H(x)$ is then a sum of homogeneous polynomials of the form (not displaying internal indices)

$$H_{m,n}(x)=[P(x)]^n E^{a_1}(x)..E^{a_m}(x) f_{m,n}[Q]_{a_1..a_m}(x)\frac{1}{[\sqrt{\det(q)}(x)]^{m+n-1}} \tag{II.2.2. 13}$$

where f is a local tensor depending only on configuration variables and their covariant derivatives with respect to A_a^i. In order to quantize (II.2.2. 13) we must point split the momenta P, E^a. Multiply (II.2.2. 13) by $1=[\frac{|\det((e_a^i))|}{\sqrt{\det(q)}}]^k$ where $k=0,1,2,...$ is an integer to be specified later on. Since up to a numerical constant $|\det((e_a^i))|$ equals $\epsilon^{abc}\epsilon_{ijk}\{A_a^i,V(R)\}\{A_b^j,V(R)\}\{A_c^k,V(R)\}$ for some appropriately chosen region we see that this factor is worth Dk volume functionals in the numerator and k factors of $\sqrt{\det(q)}$ in the denominator. We now introduce $m+n+k-1$ point splittings by the point-splitting functions $\chi_{\epsilon,x}(y)/\epsilon^D$ of the previous subsection to point split both the momenta and the factors of $|\det((e_a^i))|$. The factor $1/\epsilon^{D(m+n+k-1)}$ can be absorbed into the $\sqrt{\det(q)}$'s as before so that we get a power of $m+n+k-1$ of volume functionals of the form $V(U_\epsilon(x))$ in the denominator. Now choose k large enough until $Dk>m+n+k-1$ or $(D-1)k>m+n-1$. By suitably choosing the arguments in the process of point-splitting and choosing $R_.=U_\epsilon(.)$ we can arrange, as in the previous subsection, that the only dependence of (II.2.2. 13) on the volume functional is through Dk factors of the form

$$\frac{\{A_a^i,V(U_\epsilon)\}}{V(U_\epsilon)^{\frac{m+n+k-1}{Dk}}}=\frac{\{A_a^i,V(U_\epsilon)^{1-\frac{m+n+k-1}{Dk}}\}}{1-\frac{m+n+k-1}{Dk}} \tag{II.2.2. 14}$$

so that the volume functional is removed from the denominator. The rest of the quantization proceeds by choosing a triangulation of σ replacing connections by holonomies along its edges, Higgs fields by point holonomies at vertices, momenta by functional derivatives and Poisson brackets by commutators. By carefully choosing the factor ordering (momenta to the right hand side) one always finds a densely defined operator whose limit (as the regulator is removed) exists in the URST and whose commutator algebra is non-anomalous.

The proof shows that the density weight of one for $H(x)$ was crucial : If it would be lower than one then point splitting would result in a regulated operator whose limit is the zero operator and if it is higher than one then the limit diverges as said already earlier. Notice that the final result suffers from factor ordering *ambiguities* but *not* from factor ordering *singularities*.

II.3 Semiclassical Analysis

Despite the positive mathematical results concerning the quantization of $H(N)$, there are good reasons to be at least careful about accepting that it is physically correct. There are a number of reasons for this:

i) *Commutator Algebra*
We have seen that the commutator algebra of the Hamiltonian constraints among each other on $\mathcal{H}^0$ does not obviously resemble the classical Poisson bracket algebra. One possible reaction would be: "I could not care less[3] !" The reason is that all that is physically important is that the constraint algebra be represented correctly on the physical Hilbert space. Let us give an example: Suppose we have a classical Poisson bracket algebra of functions $J_j = \epsilon_{jkl} x_k p_l$ on the phase space $T^*\mathbb{R}^3$ given by $\{J_j, J_k\} = \epsilon_{jkl} J_l$ and that we would like to impose the constraints $\hat{J}_j \psi = 0$. Certainly we can quantize $\hat{J}_j = \epsilon_{jkl} \hat{x}_k \hat{p}_l$ and obtain a representation of the Poisson bracket algebra on the kinematical Hilbert space $\mathcal{H}^0 = L_2(\mathbb{R}^3, d^3x)$ in the usual fashion. Consider now instead the representation $\hat{J}_1' = \partial/\partial\theta$, $\hat{J}_2' = \partial/\partial\varphi$, $\hat{J}_3' = 0$ where θ, φ are the usual angular coordinates. This choice is motivated by the fact that the $\hat{J}_k$ are linear combinations of the $\hat{J}_k'$. Now, although the algebra of the $\hat{J}_k'$ is Abelean on $\mathcal{H}^0$ both sets of constraints select the same space of solutions, namely the wave functions that depend on the radial coordinate only.

The example shows that one could be lucky, that is, the quantum evolution of unphysical states may not resemble the classical evolution of unphysical functions while there is a match for physical states and physical observables, however, it would be far more convincing in our case to have a match at the kinematical level as well because *we do not know what the physical observables of the theory are.*

ii) *Non-Perturbative Hilbert Space*
The kinematical Hilbert space $\mathcal{H}^0$ is a non-perturbative one which is drastically different from the usual perturbative Fock spaces of quantum field theory on a given background spacetime. Therefore, all the beautiful coherent state machinery that is available for Fock spaces and the associated intuition that quantum field theorists have developed over the last seven decades is completely lost. That a given operator has the correct classical limit is no longer "obvious by inspection".

iii) *Tremendous Non-Linearity*
Not only is $H(N)$ is no polynomial in the basic variables A, E, it is not even an analytic function of those. The way we defined it involved the volume functional which itself is not an analytic function and the replacement of its Poisson brackets by commutators. Next, the basic operators, due to background independence are not smeared in the standard way. Finally, we had to invoke a choice function. All these steps are *never performed in standard quantum field theory* since one is always dealing with polynomials and can drop the choice function from the beginning because the available background metric fixes the (point splitting) regularization in a unique way.

Thus, what we need are new tools in order to investigate what the (semi)classical limit of the theory is and to get control on it, it is in fact the next logical step in the quantization programme. In particular, we want to verify that the Hamiltonian constraint operator really has the classical

[3] Comment by Bryce DeWitt on a talk by the author during the Meeting "MG IX", Rome, July 2000.

Hamiltonian constraint as its classical limit or to find out in the course of the analysis whether it must be modified and how. It is therefore necessary to understand in general what one means by the classical limit of the background independent quantum field theory that we have at our disposal now.

Roughly, the idea is to construct states with respect to which the gravitational degrees of freedom behave almost classical, that is, their fluctuations are minimal. It is clear that in order to test the correctness of any constraint operator one has to construct first of all kinematical semiclassical states since one cannot test an operator on its kernel. After we have made sure that the quantization of the Hamiltonian constraint is admissable, we will pass to physical coherent states.

Three proposals for semiclassical states have appeared in the literature so far:
Historically the first ones are the so-called "geometric weaves" [201] which try to approximate kinematical geometric operators only. Also "connection weaves" have been considered [202] (see also [203] for a related proposal) which are geared to approximate kinematical holonomy operators. Finally, one can get rid of a certain graph dependence of geometrical weaves through a clever statistical average [204] resulting in "statistical weaves".

The second proposal is based on the construction of coherent states for full nonlinear, non-Abelean Quantum General Relativity [207, 208, 209, 210] with all the desired properties like overcompleteness, saturation of the Heisenberg uncertainty relation, peakedness in *phase space* (thus both connection and electric flux are well approximated), construction of annihilation and creation operators and corresponding Ehrenfest theorems. Given such a coherent state, its excitations can be interpreted as the analogue of the usual graviton states [211]. One can combine these methods with a statistical average of the kind considered above to elimintae the graph dependence. The states are naturally cylindrical projections of distributions in $C^\infty(\overline{\mathcal{A}})^*$.

Finally, the third proposal [212] seems to be especially well suited for the semi-classcal analysis for free Maxwell theory and linearized gravity. It uses a striking isomorphism between the the usual Poisson algebra in terms of connections smeared in D dimensions and unsmeared electric fields on the one hand and the algebra obtained by one-dimensionally smeared connections and electric fields smeared in D dimensions on the other hand. Using this observation, which however does not carry over to the non-Abelean case, one can carry Fock like coherent states into distributions over $C^\infty(\overline{\mathcal{A}})$ and drag the Fock inner product into an inner product on the space of these distributions. See also [213] for closely related work. In [214] it is shown that, for the Abelean case, the dragged Fock measure and the uniform measure are mutually singular with respect to each other and that the dragged Fock measure does not support an electric field operator smeared in $D-1$ dimensions which are essential to use in the non-Abelean case. This indicates that all the nice structure that comes with $U(1)$ does not generalize to $SU(2)$. Nontheless the formula for these distributions suggests a transcription to the non-Abelean case [215] but it remains to be seen whether the non-distributional cylindrical projections of these distributional Fock states (called "shadows" there) have the desired semiclassical properties.

In what follows we will describe these proposals in some detail, however, since many details are still in flow we will restrict to presenting the main ideas without going too much into the technicalities.

II.3.1 Weaves

a) *Geometrical Weaves*
The early geometric weaves (first reference in [201]) were constructed as follows:
Let q^0_{ab} be a background metric. Notice that we are not introducing some background dependence here, all states still belong to the background independent Hilbert space $\mathcal{H}^0$, we are just

looking for states that have low fluctuations around a given classical three – metric. Using that metric, sprinkle non-intersecting (but possibly linked), circular, smooth loops at random with mean separation ϵ and mean radius ϵ (as measured by q^0_{ab}. The union of these loops is a graph, more precisely a link γ without intersections. The used random process was, however, not specified in [201]. Consider the state given by the product of the traces of the holonomies along those loops. The reason for choosing non-intersecting loops was that such a state was formally annihilated by the Hamiltonian constraint. Consider any surface S. From our discussion in section I.4 it is clear that this state is an eigenstate of the area operator $\widehat{\mathrm{Ar}}(S)$ with eigenvalue $\ell_p^2\sqrt{3}N(S,q^0,\epsilon)/4$ where $N(S,q^0,\epsilon)$ is the number of intersections of S with the link γ. If q^0 does not vary too much at the scale ϵ then this number is roughly given by $\mathrm{Ar}_{q^0}(S)/\epsilon^2$. Notice that all of this was done still in the complex connection representation and therefore outside of a Hilbert space context. Yet, the eigenvalue equation $\ell_p^2\mathrm{Ar}_{q^0}(S)/\epsilon^2$ tells us *that canonical quantum gravity seems to have a built in finiteness*: It does not make sense to take an arbitrarily fine graph $\epsilon \to 0$ since the eigenvalue would blow up. In order to get the corrrect eigenvalue one must take $\epsilon \approx \ell_p$, that is, the loops have to be sprinkled at Planck scale separation. This observation rests crucially on the fact that there is an area gap.

These calculations were done for metrics q^0 that are close to being flat. In the second reference of [201] weaves for Schwarzschild backgrounds were considered and require an adaption of the sprinkling process to the local curvature of q^0 in order that one obtains reasonable results.

Finally, in the third reference of [201] the link γ was generalized to disjoint collections of triples of smooth multi – loops. Each triple intersects in one point with linearly independent tangents there. The motivation for this generalization was that then the volume operator (which vanishes if there are no intersections) could also be approximated by the same technique.

b) *Connection Weaves*
For an element h of $SU(2)$ we have $\mathrm{Tr}(h) \le 2$ where equality is reached only for $h = 1$. Thus $h \mapsto 2 - \mathrm{tr}(h)$ is a non-negative function. Let now α be one of the loops considered in the third reference of [201] and let $A \in \overline{\mathcal{A}}$. Then $A \mapsto e^{-\beta[2-\mathrm{tr}(A(\alpha))]}$ is sharply peaked at those $A \in \overline{\mathcal{A}}$ with $A(\alpha) = 0$, that is, at a flat connection (since the α are contractible). Arnsdorf [202] then considers the product of all those functions which is concentrated on those distributional connections which, when restricted to the subgroupoid $l = l(\gamma)$, are flat (this function is precisely of the form of the exponential of the Wilson action employed in lattice gauge theory [176]).

Since [202] is written in the context of the Hilbert space $\mathcal{H}^0$ and since non-compact topologies of σ were considered, in contrast to [201] one had to deal with the case that the graph γ becomes infinite (the number of loops becomes infinite). Since such a state is not an element of $\mathcal{H}^0$, Arnsdorf constructed a positive linear functional on the algebra of local operators using that formal state and then used the GNS construction (see section III.6) in order to obtain a new Hilbert space in which one can now compute expectation values of various operators. Expectedly, holonomy operators along paths in l have expectation values close to their classical value at flat connections while the semi-classical behaviour of electric flux operators is less clear.

c) *Statistical Weaves*
In both the geometric and connection weave construction an arbitrary but fixed graph γ had

to be singled out. This is unsatisfactory because it involves a huge amount of arbitrariness. Which graph should one take ? Also, unless the graph γ is sufficiently random the expectation values, say of the area operator in a geometric weave for a flat background metric q^0 is not rotationally invariant.

To improve this, Bombelli [204] has employed the Dirichlet – Voronoi construction, often used in statistical mechanics [205], to the geometrical weave. Roughly, this works as follows:
Given a background metric q^0, a *compact* hypersurface σ, and a density parameter λ one can construct a subset $\Gamma(q^0,\lambda) \subset \Gamma^\omega_0$ of piecewise analytic graphs each of which, in D spatial dimensions, is such that each of its vertices is $(D+1)-$valent. A member $\gamma_{x_1,..,x_N} \in \Gamma(q^0,\lambda)$ is labelled by $N \approx [\lambda \mathrm{Vol}_{q^0}(\sigma)]$ points $x_k \in \Sigma$ where [.] denotes the Gauss bracket. The graph $\gamma_{x_1,..,x_N}$ is obtained unambiguously from the set of points $x_1,..,x_N$ and the metric q^0 (provided that it is close to being flat) by employing natural notions like minimal geodesic distances etc. Next, given a spin label j and an intertwiner I we can construct a gauge invariant spin-net $s_{x_1,..,x_N}(j,I)$ by colouring each edge with the same spin and each vertex with the same intertwiner. From these data one can construct the "density operator"

$$\hat{\rho}(q^0,\lambda,I,j) := \int_{\sigma^N} d\mu_{q^0}(x_1)..d\mu_{q^0}(x_N)\; T_{s_{x_1,..,x_N}(j,I)} < T_{s_{x_1,..,x_N}(j,I)}, . > \tag{II.3.1. 1}$$

where

$$d\mu_{q^0}(x) := \frac{\sqrt{\det(q^0)(x)}d^D x}{\mathrm{Vol}_{q^0}(\sigma)} \tag{II.3.1. 2}$$

is a probability measure (it is here where compactness of σ is important). The reason for the inverted commas in "density operator" is that (II.3.1. 1) actually is the zero operator [206]. To see this, notice that for any spin-network state T_s we have $< T_{s_{x_1,..,x_N}(j,I)}, T_s >= \delta_{s_{x_1,..,x_N}(j,I),s}$ which in particular means that $\gamma_{x_1,..,x_N} = \gamma(s)$. But the set of points satisfying this is certainly thin with respect to the measure (II.3.1. 2). What happens is that although for any spin-network state T_s the one-dimensional projector $T_s < T_s. >$ is a trace class operator of unit trace, the trace operation does not commute with the integration in (II.3.1. 1). However, one can then define a positive linear functional $\omega_{q^0,\lambda,I,j}$ on the algebra of linear operators on $\mathcal{H}^0$ by

$$\omega_{q^0,\lambda,j,I}(\hat{O}) := \int_{\sigma^N} d\mu_{q^0}(x_1)..d\mu_{q^0}(x_N)\;\; < T_{s_{x_1,..,x_N}(j,I)}, \hat{O} T_{s_{x_1,..,x_N}(j,I)} > \tag{II.3.1. 3}$$

which would equal $\mathrm{Tr}(\hat{\rho}(q^0,\lambda,j,I)\hat{O})$ if integration and trace would commute. Via the GNS construction one can now define a new representation $\mathcal{H}^0_{q^0,\lambda,j,I}$ which now depends on a background structure. The representations $\mathcal{H}^0$ and $\mathcal{H}^0_{q^0,\lambda,j,I}$ are certainly not comparable in the sense that one can embed one space into the other and presumably they are (unitarily) inequivalent.

What is interesting about (II.3.1. 3) is that for an exactly flat background the expectation values of, say the area operator, are *Euclidean invariant*. In order to match the expectation values of $\widehat{\mathrm{Ar}}(S)$ with the value $\mathrm{Ar}_{q^0}(S)$ one must choose j according to $[\sqrt{j(j+1)}\ell_p^2\beta\lambda^{2/3}/2] = 1$. A similar calculation for the volume operator presumably fixes the value I for the intertwiner.

II.3.2 Coherent States

Especially the statistical weave construction of the previous subsection looks like a promising starting point for semiclassical analysis. However, there are several drawbacks with weaves:

i) *Phase Space Approximation*
All the weaves discussed above seem to approximate either the connection or the electric field appropriately although the degree of their approximation has never been checked (are the fluctuations small ?). However, what we really need are states which approximate the connection and the electric field simultaneously with small fluctuations.

ii) *Arbitrariness of Spins and Intertwiners*
All weaves proposed somehow seem to arbitrarily single out special and uniform values for spin and intertwiners. Drawing an anology with a system of uncoupled harmonic oscillators, it is like trying to build a semiclassical state by choosing an arbitrary but fixed occupation number (spin) for each mode (edge). However, we know that the preferred semiclassical states for the harmonic oscillator are coherent states which depend on all possible occupation numbers. As we will see, issue i) and ii) are closely related.

iii) *Arbitrariness of Graphs*
Even in the statistical weave construction we select arbitrarily only a certain subclass of graphs. Again, drawing an anology with the harmonic oscillator picture, this is like selecting a certain subset of modes in order to build a semiclassical state. However, then not all modes can behave semi-classically.

iv) *Missing Construction Principle*
The weave states constructed suffer from a missing enveloping construction principle that would guarantee from the outset that they possess desired semi-classical properties.

The aim of the series of papers [207, 208, 209, 210] was to decrease this high level of arbitrariness, to look for a systematic construction principle and to make semiclassical states for quantum gravity look more similar to the semiclassical states for free Maxwell theory which are in fact *coherent states* and have been extremely successful, see e.g. [216] and referencese therein.

II.3.2.1 Semiclassical States and Coherent States

Recall that quantization is, roughly speaking, an attempt to construct a * homomorphism

$$\bigwedge : \; (\mathcal{M}, \{.,.\}, \mathcal{O}, \overline{(.)}) \to (\mathcal{H}, \frac{[.,.]}{i\hbar}, \widehat{\mathcal{O}}, (.)^\dagger) \tag{II.3.2. 1}$$

from a subalgebra $\mathcal{O} \subset C^\infty(\mathcal{M})$ of the Poisson algebra of complex valued functions on the symplectic manifold $(\mathcal{M}, \{.,.\})$ to a subalgebra $\widehat{\mathcal{O}} \subset \mathcal{L}(\mathcal{H})$ of the algebra of linear operators on a Hilbert space $\mathcal{H}$ with inner product $< .,. >$ such that Poisson brackets turn into commutators and complex conjugation into the adjoint operation. Notice that the map cannot be extended to all of $C^\infty(\mathcal{M})$ (only up to quantum corrections) unless one dives into deformation quantization, see e.g. [217] and references therein, the subalgebra for which it holds is referred to as the algebra of elementary functions (operators). The algebra $\mathcal{O}$ should be sufficiently large in order that more complicated functions can be expressed in terms of elements of it so that they can be quantized by choosing a suitable factor ordering.

Dequantization is the inverse of the map (II.3.2. 1). A possible way to phrase this more precisely is:

Definition II.3.1
A system of states $\{\psi_m\}_{m\in\mathcal{M}} \in \mathcal{H}$ is said to be semiclassical for an operator subalgebra $\overline{\mathcal{O}} \subset \mathcal{L}(\mathcal{H})$ provided that for any $\hat{O}, \hat{O}' \in \widehat{\mathcal{O}}$ and any generic point $m \in \mathcal{M}$

[1.] Expectation Value Property

$$|\frac{<\psi_m, \hat{O}\psi_m>}{O(m)} - 1| \ll 1 \tag{II.3.2. 2}$$

[2.] Infinitesimal Ehrenfest Property

$$|\frac{<\psi_m, [\hat{O}, \hat{O}]\psi_m>}{i\hbar\{O, O'\}(m)} - 1| \ll 1 \tag{II.3.2. 3}$$

[3.] Small Fluctuation Property

$$|\frac{<\psi_m, \hat{O}^2\psi_m>}{<\psi_m, \hat{O}\psi_m>^2} - 1| \ll 1 \tag{II.3.2. 4}$$

The quadruple $(\mathcal{M}, \{.,.\}, \mathcal{O}, \overline{(.)})$ is then called the classical limit of $(\mathcal{H}, \frac{[.,.]}{i\hbar}, \widehat{\mathcal{O}}, (.)^\dagger)$.

Clearly definition II.3.1 makes sense only when none of the denominators displayed vanish so they will hold at most at generic points m of the phase space (meaning a subset of $\mathcal{M}$ whose complement has Liouville measure comparable to a phase cell) which will be good enough for all practical applications. Notice that if [1.] holds for $\hat{O}$ then it holds for $\hat{O}^\dagger$ automatically. Condition [1.] is for polynomial operators sometimes required in the stronger form that (II.3.2. 2) should vanish exactly which can always be achieved by suitable (normal) ordering prescriptions. Condition [2.] ties the commutator to the Poisson bracket and makes sure that the infinitesimal quantum dynamics mirrors the infinitesimal classical dynamics. If the error in [2.] vanishes then we have a finite Ehrenfest property which in non-linear systems is very hard to achieve. Finally, [3.] controls the quantum error, the fluctuation of the operator.

Coherent states have further properties which can be phrased roughly as follows:

Definition II.3.2
A system of states $\{\psi_m\}_{m\in\mathcal{M}} \in \mathcal{H}$ is said to be coherent for an operator subalgebra $\hat{\mathcal{O}} \subset \mathcal{L}(\mathcal{H})$ provided that for any $\hat{O}, \hat{O}' \in \widehat{\mathcal{O}}$ and any generic point $m \in \mathcal{M}$ in addition to properties [1.], [2.] and [3.] we have

[4.] Overcompleteness Property
There is a resoltion of unity

$$1_{\mathcal{H}} = \int_{\mathcal{M}} d\nu(m)\psi_m <\psi, .> \tag{II.3.2. 5}$$

for some measure ν on $\mathcal{M}$.

[5.] Annihilation Operator Property
There exist elementary operators $\hat{g}$ (forming a complete system) such that

$$\hat{g}\psi_m = g(m)\psi_m \tag{II.3.2. 6}$$

[6.] Minimal Uncertainty Property
For the self-adjoint operators $\hat{x} := (\hat{g} + \hat{g}^\dagger)/2,\ \hat{y} := (\hat{g} - \hat{g}^\dagger)/(2i)$ the (unquenched) Heisenberg uncertainty relation is saturated

$$<(\hat{x}- <\hat{x}>_m)^2>_m = <(\hat{y}- <\hat{y}>_m)^2>_m = \frac{1}{2}| <[\hat{x}, \hat{y}]>_m | \tag{II.3.2. 7}$$

[7.] Peakedness Property
For any $m \in \mathcal{M}$, the overlap function

$$m' \mapsto |<\psi_m, \psi_{m'}>|^2 \tag{II.3.2. 8}$$

is concentrated in a phase cell of Liouville volume $\frac{1}{2}|<[\hat{p},\hat{h}]>_m|$ if $\hat{p}$ is a momentum operator and $\hat{h}$ a configuration operator.

These four conditions are not completely independent of each other, in particular, [5.] implies [6.] but altogether [1.] – [7.] comprises a fairly complete list of desirable properties for semiclassical (coherent states).

II.3.2.2 Construction Principle: Complexifier Method and Heat Kernels

Usually one introduces coherent states for the harmonic oscillator as eigenstates of the annihilation operator in terms of superpositions of energy eigenstates. This method has the disadvantage that one needs a preferred Hamiltonian, that is, dynamical input in order to define suitable annihilation operators. Even if one has a Hamiltonian, the construction of annihilation operators is no longer straightforward if we are dealing with a non-linear system. Since we neither have a Hamiltonian nor a linear system and since for the time being we are anyway interested in kinematical coherent states, we have to look for a different constructive strategy.

A hint comes from a different avenue towards the harmonic oscillator coherent states. Let the Hamiltonian be given by

$$H := \frac{1}{2}[p^2/m + m\omega^2 X^2] = \omega\bar{z}z \text{ where } z = \frac{\sqrt{m\omega}x - ip/\sqrt{m\omega}}{\sqrt{2}} \tag{II.3.2. 9}$$

Define the *complexifier* function

$$C := \frac{p^2}{2m\omega} \tag{II.3.2. 10}$$

then it is easy to see that

$$z = \sqrt{\frac{m\omega}{2}} \sum_{n=0}^{\infty} \frac{(-i)^n}{n!}\{C, x\}_n \tag{II.3.2. 11}$$

(recall that in our terminology $\{p, x\} = 1$). Translating this equation into quantum theory we find

$$\hat{z} = \frac{\sqrt{m\omega}}{\sqrt{2}} \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \frac{[\hat{C}, \hat{x}]_n}{(i\hbar)^n} = e^{-t(-\Delta/2)} \frac{\hat{x}\sqrt{m\omega}}{\sqrt{2}} (e^{-t(-\Delta/2)})^{-1} \tag{II.3.2. 12}$$

where the *classicality parameter*

$$t := \hbar/(m\omega) \tag{II.3.2. 13}$$

has naturally appeared and which for this system has dimension cm^2. The operator $\hat{z}$ is usually chosen by hand as the annihilation operator. Let us define that coherent states ψ_z are eigenstates of $\hat{z}$. Given formula (II.3.2. 13) we can trivially construct them as follows: Let δ_x be the $\delta-$distribution, supported at x, with respect to the Hilbert space measure dx. Define $\psi_x := e^{-t\hat{C}/\hbar^2}\delta_x$. Then formally

$$\hat{z}\psi_x = e^{-t\hat{C}/\hbar^2} \frac{\sqrt{m\omega}\hat{x}}{\sqrt{2}}\delta_x = \frac{x\sqrt{m\omega}}{\sqrt{2}}\psi_x \tag{II.3.2. 14}$$

because δ_x is an eigendistribution of the operator $\hat{x}$. The crucial point is now that ψ_x is an analytic function of x as one can see by using the Fourier representation for the $\delta-$distribution $\delta_x = \int_{\mathbb{R}} dk/(2\pi)e^{ikx}$. We can therefore analytically extend ψ_x to the complex plane $x \to x - ip/(m\omega)$ and arrive with the trivial redefinition $\psi_{x-ip/(m\omega)} \mapsto \psi_z$ at

$$\hat{z}\psi_z = z\psi_z \tag{II.3.2. 15}$$

One can check that the state $\psi_z/||\psi_z||$ coincides with the usual harmonic oscillator coherent states up to a phase.

We see that the harmonic oscillator coherent states can be naturally put into the language of the Wick rotation transform of section II.1.1. This observation, stripping off the particulars of the harmonic oscillator, admits a generalization that applies to any symplectic manifold $\mathcal{M}, \{.,.\}$ which is a cotangent bundle $\mathcal{M} = T^*\mathcal{C}$ where $\mathcal{C}$ is the configuration base space of $\mathcal{M}$. The essential steps can be summarized in the following algorithm (we suppress all indices, discrete and continuous):

1) *Hilbert Space and $\delta-$Distribution*
The Hilbert space is supposed to be an L_2 space, that is, there exists a measure μ on $\mathcal{C}$ such that $\mathcal{H} = L_2(\mathcal{C}, d\mu)$. With respect to the measure μ we may define the $\delta-$distribution supported at $x \in \mathcal{C}$ by the formula $\delta_x(f) := \int_{\mathcal{C}} d\mu(x')\delta_x(x')f(x') = f(x)$ for any $f \in C_0^\infty(\mathcal{C})$ (or any other dense space of tests functions). Here we have denoted the integral kernel of the distribution by $\delta_x(x')$.

2) *Complexifier and Heat Kernel Evolution*
Find a non-negative function C on $\mathcal{M}$ which can be quantized on $\mathcal{H}$ as a positive definite, self-adjoint operator. Moreover, the dimensions of $x, \{C, x\}$ should coincide. Then $e^{-\hat{C}/\hbar}$ is a bounded operator and can be defined via the spectral theorem. Furthermore, we need that the heat kernel evolution of the $\delta-$distribution $\psi_x := e^{-\hat{C}/\hbar}\delta_x$ is a square integrable function in $\mathcal{H}$ which at the same time is analytic in x.

3) *Analytic Continuation and Annihilation Operators*
Let ψ_z be the analytic continuation of ψ_x and define the anniliation operator $\hat{z} := e^{-\hat{C}/\hbar}\hat{x}(e^{-\hat{C}/\hbar})^{-1}$. Then automatically $\hat{z}\psi_z = z\psi_z$ is an eigenstate. Notice that the inverse $(e^{-\hat{C}/\hbar})^{-1}$ is only densely defined on functions of the form $e^{-\hat{C}/\hbar}f,\ f \in \mathcal{H}$ which have been smoothened out by $e^{-\hat{C}/\hbar}$.

4) *Classicality Parameter and Physical Interpretation*
The quantity $\hat{C}/\hbar$ is dimensionfree by construction. The classicality parameter t is defined by $\hat{C}/\hbar = -t\Delta/2$ where Δ is a negative definite differential operator of order greater than one in order that, according to the rule $\hat{p} = i\hbar\partial/\partial x$, the parameter t is proportional to a positive power of $\hbar$ and therefore small. It is clear that $\hat{z}$ is a quantization of $z := \sum_{n=0}^\infty \frac{(-i)^n}{n!}\{C, x\}_n$. We require further that C has been chosen in such a way that the functions $z, \bar{z}$ suffice to isolate configuration and momentum function x, p respectively. Thus we have an invertible map $m = (x, p) \mapsto (z, \bar{z})$ and can finally define $\psi_m := \psi_{z(m)}$.

Certainly, steps 1) – 4) are only formal and have to be justified mathematically in the model at hand. However, given an L_2 Hilbert space over the configuration space, they merely require *one input*: the choice of the complexifier C and which one selects depends on some physical input.

The complexifier method is extremely natural: Besides the fact that, as one can show, any coherent states that have been constructed for linear field theories actually fall into the catgory of states that have been constructed by the complexifier method, automatically the following coherent state properties (formally) hold: The annihilation operator property [5.] trivially holds by construction and hence the saturation of the unquenched minimal uncertainty relation, property [6.], as well. Moreover, the expectation value property [1.] automatically holds for any normal ordered polynomial of the $\hat{z}, \hat{z}^\dagger$. The overercompleteness property [4.] is equivalent to showing that the coherent state transform $(\hat{U}f)(z) := [e^{-\hat{C}/\hbar}f(x)]_{|x\to z}$ introduced in section II.1.1 is unitary and we have given there a formal recipe that constructs a measure ν on the complexification $\mathcal{C}^{\mathbb{C}}$ such that the transform becomes a partial isometry at least (the hard part is to show that the transform is onto the space of holomorphic $\nu-$square integrable functions on $\mathcal{C}^{\mathbb{C}}$). The peakedness property [7.] is at least rather likely to hold because what $e^{-\hat{C}/\hbar}$ does to the $\delta-$distribution (which is *sharply peaked*) is to decrease the size of the peak and to increase its width (of the order $\sqrt{t}$) at least in the configuration representation. Next, again for polynomials of $\hat{z}, \hat{z}^\dagger$ the infinitesimal Ehrenfest property [2.] should follow from the correct quantization of $\hat{p}, \hat{x}$ (less trivial are non-polynomial functions, which however crucially appear in our applications). Finally, the small fluctuation property [3.] trivially holds for polynomials of $\hat{z}$ alone or $\hat{z}^\dagger$ alone and therefore holds for more general polynomials as well if [2.] holds.

This concludes our motivation for considering complexifier coherent states.

II.3.2.3 Coherent States for Canonical Quantum General Relativity

Let us now apply the framework of the previous subsection to canonical quantum general relativity. We have a (quantum) configuration space $\mathcal{C} = \overline{\mathcal{A}}$ and a measure μ_0 thereon which together build the Hilbert space $\mathcal{H}^0 = L_2(\overline{\mathcal{A}}, d\mu_0)$. The $\delta-$distribution on $\overline{\mathcal{A}}$ with respect to the measure μ_0 is given by

$$\delta_A = \sum_{s\in\mathcal{S}} T_s(A) < T_s, . > \qquad \text{(II.3.2. 16)}$$

So in principle, all that remains to do is to find a suitable complexifier. A natural choice is the volume operator $\hat{\mathrm{Vol}}(\sigma)$ as the complexifier [207] because it is background independent, gauge invariant, spatially diffeomorphism invariant, a differential operator of order $3/2 > 1$, positive semi-definite and selfadjoint. In order that classically $A, \{C, A\}$ have the same dimension we will choose a parameter a with dimension of length whose physical significance will become clear only later and define $C := 2\mathrm{Vol}(\sigma)/(\kappa a)$ (again we take $\beta = 1$). Then

$$\hat{C}/\hbar = \frac{\ell_p}{a}\frac{\hat{V}(\sigma)}{\ell_p^3} \qquad \text{(II.3.2. 17)}$$

is dimensionfree and the classicality parameter is given by $t = \ell_p/a$ which should be much smaller than unity.

It is easy to see that for any path $p \in \mathcal{P}$ we have

$$\sum_{n=0}^{\infty} \frac{(-i)^n}{n!}\{C, h_p\}(A, E) = h_p(A - ie/a) \qquad \text{(II.3.2. 18)}$$

where $e = e_a^j dx^a \tau_j/2$ is the co-triad one-form, that is, we get the holonomy of an $SL(2, \mathbb{C})$ connection. Therefore, we know the classical correspondence of the annihilation operators

$$\hat{g}_p := e^{-\hat{C}/\hbar}\hat{h}_p e^{\hat{C}/\hbar} \qquad \text{(II.3.2. 19)}$$

The crucial question is now whether $e^{-\hat{C}/\hbar}\delta_A$ is analytic in A. In order to compute this quantity we switch to a new orthonormal basis $s=(\gamma(s),\lambda(s),I(s))$ where for given γ we have diagonalized all spin-network states over γ with eigenvalues λ of $\hat{V}(\sigma)/\ell_p^3$ and degeneracy labels I. Then

$$e^{-\hat{C}/\hbar}\delta_A=\sum_s e^{-t\lambda(s)}T_s(A)<T_s,.> \tag{II.3.2. 20}$$

Since the functions $T_s(A)$ are analytic in A we may define our coherent states to be

$$\psi_{A^{\mathbb{C}}}:=\sum_s e^{-t\lambda(s)}T_s(A^{\mathbb{C}})<T_s,.> \tag{II.3.2. 21}$$

where $A^{\mathbb{C}}\in\overline{\mathcal{A}}^{\mathbb{C}}:=\mathrm{Hom}(\mathcal{L},SL(2,\mathbb{C}))$. Of course, (II.3.2. 21) can be applied to any suitable complexifier C.

There are several problems with (II.3.2. 21):

1)

Although it defines an element of $\mathcal{D}^*$, it is not an element of $\mathcal{H}^0$ and in order to use these states in expectation value calculations one would need to introduce a new inner product for them just as we have to do for solutions of the Hamiltonian constraint.

There is no obvious choice for such an inner product at all. One could of course try to normalize (II.3.2. 21) but the resulting expressions become hard to control because the eigenbasis of the volume operator is not known.

2)

One could consider the "cut-off" states

$$\psi_{\gamma,A^{\mathbb{C}}}:=\sum_{\gamma'\subset\gamma}\;\sum_{s;\,\gamma(s)=\gamma'} e^{-t\lambda(s)}T_s(A^{\mathbb{C}})<T_s,.> \tag{II.3.2. 22}$$

where the sum is over all subgraphs of γ, that is, those that arise by removing edges from $E(\gamma)$, one by one in all possible ways. (Here a difference arises depending on whether we work at the gauge variant or gauge invariant level because in the latter case one can remove edges only in such a way that γ' has no univalent vertices. Also if γ contains three edges e_1,e_2,e_3 which meet in one point and such that $e_1\circ e_2$ is analytic, then after removing e_3 we take the convention that the point $e_1\cap e_2$ is still a vertex of $\gamma-\{e_3\}$ in the gauge variant case). But even for those it is neither clear how to calculate anything nor is it clear whether (II.3.2. 22) is an element of $\mathcal{H}^0$ at all because the degeneracy N_λ of almost all λ on any given graph could exceed the damping factor $e^{-t\lambda}$. So again the complicated spectrum of the volume operator makes (II.3.2. 22) at least highly unpractical.

3)

If we would use cut-off states to do semi-classical physics, then they are presumably inadequate for computing expectation values of operators with non-vanishing matrix elements between spin network states over different graphs like the Hamiltonian – or Diffeomorphism Constraint.

4)

The Poisson algebra of the classical functions $h_p(A^{\mathbb{C}}),\overline{h_p(A^{\mathbb{C}})}$ does not close and therefore the commutators between the associated operators should look horrible, that is, the infinitesimal Ehrenfest property will be difficult to verify.

One way out is too look for a different classical function C, maybe background dependent, which at least does not have the problems 2), 3). However, *for non-Abelean gauge groups there seems to be no C, polynomial in the electric fields, such that $\hat{C}$ leaves the space of cylindrical functions invariant* *and simultaneously 2) and 3) disappear* [218] !

Another option is to construct a family of coherent states $(\psi_{\gamma,(A,E)})_{\gamma\in\Gamma}$ by hand, that is, for each γ we choose a complexifier C_γ and repeat the above procedure restricted to the Hilbert space $\mathcal{H}_\gamma$. The function C_γ should cure the problems 2) and 3) just mentioned but it is no longer required that C_γ is the discretization of some well-defined function C on $\mathcal{M}$, in particular, it will be not be the case that the family of operators $\hat{C}_\gamma$ is consistent (although this can always be cured by defining them in the spin-network basis). This has been proposed in [207] and works as follows: Define

$$\begin{aligned}\delta_A^\gamma &:= \sum_{s\in\mathcal{S};\ \gamma(s)=\gamma} T_s(A) < T_s, . > \\ \delta_{\gamma,A} &:= \sum_{\gamma'\subset\gamma} \delta_A^{\gamma'}\end{aligned} \qquad \text{(II.3.2. 23)}$$

We evidently have the identitity

$$\delta_A = \sum_{\gamma\in\Gamma_0^\omega} \delta_A^\gamma \qquad \text{(II.3.2. 24)}$$

so that the second line in (II.3.2. 23) is the "distribution cut off at γ". A simplification arises at the gauge variant level since then evidently

$$\delta_{\gamma,A} = \prod_{e\in E(\gamma)} \delta_{e,A} \qquad \text{(II.3.2. 25)}$$

factorizes. Now $\delta_{e,A} = \delta_{A(e)}$ where the latter distribution is with respect to the Haar measure. Due to the Peter&Weyl theorem

$$\delta_h(h') = \sum_{\pi\in\Pi} d_\pi \chi_\pi(h(h')^{-1}) \qquad \text{(II.3.2. 26)}$$

which demonstrates that with $\delta_A^e = \delta_{e,a} - 1$ we also have

$$\delta_A^\gamma = \prod_{e\in E(\gamma)} \delta_A^e \qquad \text{(II.3.2. 27)}$$

Let us now specify C_γ. Given a graph γ consider a system of mutually disjoint, open surfaces $(S_e)_{e\in E(\gamma)}$ where $e\cap S_{e'} = \emptyset$ if $e \neq e'$ and $x_e := e\cap S_e$ is an interior point of both e, S_e. Moreover, S_e carries the orientation such that e is of the "up" type and the collection S_e is supposed to form a polyhedronal decomposition of σ (add some surfaces that do not intersect γ at all if necessary). Next, choose a system of non-self-intersecting paths $\rho_e(x)$ within S_e, one for every point $x\in S_e$ with $b(\rho_e(x)) = x_e$ and $f(\rho_e(x)) = x$. From these data construct the functions

$$P_j^e(A,E) := -\frac{1}{2a_e^2}\mathrm{Tr}(\tau_j \int_{S_e} \mathrm{Ad}_{A(e_{x_e}\circ\rho_e(x))}(*E(x)) \qquad \text{(II.3.2. 28)}$$

where e_{x_e} is the segment of e with $b(e_{x_e}) = b(e),\ f(e_{x_e}) = x_e$ and $*E = \epsilon_{abc}E_j^a\tau_j dx^b\wedge dx^c$. Again the length parameter a_e will receive its physical meaning only later in concrete physical applications.

The crucial fact about the system of functions h_e, P^e is that they are gauge covariant, $\lambda_g^* P^e = \mathrm{Ad}_{g(b(e))}(P^e)$, in contrast to the $E_j(S)$ of section I.3.1. 1, diffeomorphism covariant if $a_e = a$ is a constant (all edges, paths, surfaces just get mapped to diffeomorphic images) and they form a *closed*

Poisson subalgebra of $C^\infty(\mathcal{M})$ given by

$$\begin{aligned}
\{h_e, h_{e'}\} &= 0 \\
\{P_j^e, h_{e'}\} &= \frac{\kappa}{a_e^2}\delta^e_{e'}\frac{\tau_j}{2}h_e \\
\{P_j^e, P_k^{e'}\} &= -\delta^{ee'}\frac{\kappa}{a_e^2}\epsilon_{jkl}P_l^e
\end{aligned} \tag{II.3.2. 29}$$

However, this Poisson algebra is isomorphic to the natural Poisson algebra on $\mathcal{M}_\gamma := \prod_{e\in E(\gamma)} T^*(SU(2))$ so what we have achieved is construct a map

$$\Phi'_\gamma :\ \mathcal{M} \to \mathcal{M}_\gamma;\ (A, E) \mapsto (h_e(A), P_j^e(A, E))_{e\in E(\gamma)} \tag{II.3.2. 30}$$

which is a *partial symplectomorphism.* (Notice that it is neither one to one nor onto for fixed γ. Here we are abusing the notation somewhat because Φ'_γ certainly also depends on the $S_e, \rho_e(x)$). This fact is going to be fundamental for all that follows for the following reason: What we are really going to do is to construct coherent states for the phase space $M_\gamma := [T^*(SU(2))]^{|E(\gamma)|}$ and since the Poisson structures of the phase spaces $\Phi'_\gamma(\mathcal{M})$ and $\mathcal{M}_\gamma$ coincide we automatically have proved the Ehrenfest property for $\Phi'_\gamma(\mathcal{M})$. Now, if γ gets sufficiently fine, we can approximate any function on $\mathcal{M}$ by functions in $\Phi'_\gamma(\mathcal{M})$ and in that sense we are constructing *approximate coherent states* for $\mathcal{M}$.

Next we must construct C_γ. In analogy to the harmonic oscillator we choose a function which is quadratic in the momenta because this will lead to similar Gaussian peakedness properties. Thus we define

$$C_\gamma := \frac{1}{2\kappa}\sum_{e\in E(\gamma)} a_e^2 (P_j^e)^2 \tag{II.3.2. 31}$$

One may check, that this leads to the complexification

$$g_e := \sum_{n=1}^\infty \frac{(-i)^n}{n!}\{C_\gamma, h_e\}_n = e^{-iP_j^e\tau_j/2}h_e \tag{II.3.2. 32}$$

where the Poisson brackets are those of $\mathcal{M}$. Something amazing has happened in (II.3.2. 32): We have stumbled naturally on the diffeomorphism

$$T^*(SU(2)) \to SL(2,\mathbb{C});\ (h, P) \mapsto e^{-iP^j\tau_j/2}h \tag{II.3.2. 33}$$

where the inverse of (II.3.2. 33) is given by polar decomposition. Now, while the complexification of $\mathbb{R}$ is given by $\mathbb{C}$, the complexification of a Lie group G with Lie algebra Lie(G) is given by the image under the exponential map of the complexification of its Lie algebra (that is, we allow arbitrary complex coefficients θ^j of the Lie algebra basis τ_j) and (II.3.2. 32) tells us precisely how this is induced by the complexifier. The map (II.3.2. 33) allows us to identify $\mathcal{M}_\gamma$ with $SL(2,\mathbb{C})^{|E(\gamma)|}$ so that we have altogether a map

$$\Phi_\gamma :\ \mathcal{M} \to \mathcal{M}_\gamma;\ (A, E) \mapsto m_\gamma(A, E) := (g_e(A, E) := e^{-iP_j^e\tau_j/2}h_e)_{e\in E(\gamma)} \tag{II.3.2. 34}$$

The Poisson algebra (II.3.2. 29) suggests on $\mathcal{H}^0_\gamma$ the quantization $\hat{P}_j^e = it_e R_e^j/2$ while $\hat{h}_e$ is a multiplication operator. Here the classicality parameters

$$t_e := \frac{\ell_p^2}{a_e^2} \tag{II.3.2. 35}$$

have naturally appeared and it follows that

$$\hat{C}_\gamma/\hbar = -\frac{1}{2}\sum_{e\in E(\gamma)} t_e\Delta_e \tag{II.3.2. 36}$$

where $\Delta_e = (R^j_e)^2/4$. Our annihilation operators become

$$\hat{g}_e := e^{-\hat{C}_\gamma/\hbar}\hat{h}_e(e^{-\hat{C}_\gamma/\hbar})^{-1} = e^{-t_e\tau_j^2/8}e^{-i\hat{P}^e_j\tau_j/2}\hat{h}_e \tag{II.3.2. 37}$$

which up to a quantum correction is precisely the quantization of (II.3.2. 32). Then we can define abstract coherent states for $\mathcal{H}\gamma$ by

$$\begin{aligned}
\psi_{\gamma,m_\gamma} &:= [e^{-\hat{C}_\gamma/\hbar}\delta_{\gamma,h_\gamma}]_{h_\gamma\to m_\gamma} \\
&= \prod_{e\in E(\gamma)} [e^{t_e\Delta_e/2}\delta_{h_e}]_{h_e\to g_e} \\
\psi^\gamma_{m_\gamma} &:= [e^{-\hat{C}_\gamma/\hbar}\delta^\gamma_{h_\gamma}]_{h_\gamma\to m_\gamma} \\
&= \prod_{e\in E(\gamma)} [e^{t_e\Delta_e/2}\delta_{h_e}-1]_{h_e\to g_e} \\
\psi_g &:= e^{t\Delta/2}\delta_h]_{h\to g} = \sum_{j=0,1/2,1,3/2,..}(2j+1)e^{-tj(j+1)/2}\chi_j(gh^{-1})
\end{aligned} \tag{II.3.2. 38}$$

and coherent states on $\mathcal{H}^0$ by

$$\psi_{\gamma,m} := \hat{U}_\gamma\psi_{\gamma,\Phi_\gamma(m)} \text{ and } \psi^\gamma_m := \hat{U}_\gamma\psi^\gamma_{\Phi_\gamma(m)} \tag{II.3.2. 39}$$

where $\hat{U}_\gamma:\ \mathcal{H}^0_\gamma\to\mathcal{H}^0$ is the usual isometric monomorphism.

In [208] we have proved peakedness –, expectation value –, small fluctuation and Ehrenfest properties for the gauge variant states mathematical states ψ_{γ,m_γ} and the algebra of operators $\mathcal{L}(H^0_\gamma)$. All proofs can be reduced to proving it for single copy of $SU(2)$. Overcompleteness follows from the results due to Hall [79] for the states ψ_g on $L_2(SU(2),d\mu_H)$. Annihilation operators have been defined above and for those minimal uncertainty properties follow.

Next, given a system of elements $g_e\in SL(2,\mathbb{C})$, one for each analytic path $e\in\mathcal{P}$ we can form the distribution

$$\psi_g := \sum_{\gamma\in\Gamma^\omega_0} <\psi^\gamma_{\overline{g_\gamma}},.>\psi^\gamma_{g_\gamma} \tag{II.3.2. 40}$$

where $g_\gamma=\{g_e\}_{e\in E(\gamma)}$. Now as shown in [83] it is indeed possible to define an operator $\hat{C}$ through its cylindrical projections $\hat{C}_\gamma$ provided the system of positive numbers t_e satisfies the two conditions

$$t_{e_1}+t_{e_2}=t_{e_1\circ e_2} \text{ and } t_e=t_{e^{-1}} \tag{II.3.2. 41}$$

which implies that the t_e are in this case not constants. The t_e thus have all the properties of a length function and we may use the background to be approximated in order to define it. The distribution (II.3.2. 40) is then precisely of the type (II.3.2. 20). Of course, by defining $\hat{C}_\gamma$ on spin-network functions rather than cylindrical functions this can also be achieved if we do not have (II.3.2. 41).
Remark:
We could then extend the definition of the operators $\hat{g}_e$ by

$$\hat{g}_e := e^{-\hat{C}/\hbar}\hat{h}_e(e^{-\hat{C}/\hbar})^{-1} \tag{II.3.2. 42}$$

from which the properties $\hat{g}_e \hat{g}_{e'} = \hat{g}_{e \circ e'}$ and $\hat{g}_{e^{-1}} = \hat{g}_e^{-1}$ due to similar properties for the operator $\hat{h}_e$. It follows that if the label g in (II.3.2. 40) is to reproduce all the properties of the operators $\hat{g}_e$ then we should have $g_e g_{e'} = g_{e \circ e'}$, $g_{e^{-1}} = g_e^{-1}$ in other words, g qualifies as a generalized connection, that is, an element of $\mathrm{Hom}(\mathcal{P}, SL(2,\mathbb{C}))$. The question is then whether the images $\Phi_e(m)$ defined in (II.3.2. 30) do have those properties for all $m \in \mathcal{M}$ by choosing the $S_e, \rho_e(x)$ appropriately. Notice that for the volume operator as the complexifier this property would be trivially satisfied but in our case the answer is less clear.

In any case, the purpose of (II.3.2. 40) is to demonstrate that with our family of coherent states it is possible to form a distribution which is graph independent but unfortunately one does not have an inner product on these objects available and thus the best thing that one can do at this point is to take the cut-off states $\psi_{\gamma,m}$ or ψ^{γ}_m. Notice that if for $\gamma' \subset \gamma$ we choose the $S_e, \rho_e(x)$; $e \in E(\gamma')$ to be those that we chose for γ then $\psi_{\gamma,m} = \sum_{\gamma' \subset \gamma} \psi^{\gamma}_m$. The coherent state properties that we established hold for the $\psi_{\gamma,m}$ and to some extent also for the ψ^{γ}_m.

We then must deal with the question of how to choose γ. This question is analyzed in detail in [210]. One possibility is to form a density matrix similar to the one we discussed above but averaging only over a countable number of states (thus not leaving $\mathcal{H}^0$). Another would be to choose for γ a generic random graph which does not display any direction dependence on large scales. In any of these scenarios the picture that arises is the following:
Given γ, m we can extract from these two data two scales: The first is a *graph scale* ϵ given by the average edge length as measured by m. The second is a *curvature scale* L which is determined both by the mean curvature radius of the four-dimensional metric determined by m and the mean curvature of the induced metric on the embedded submanifolds $e, S_e, \rho_e(x)$ (so that even in the case that m are exactly flat initial data the scale L is not necessarily infinity). We then must decide which (kinematical) observables should behave maximally semi-classically. This is a choice that must be made and the choice of γ will depend largely on this physical input. In [210] we chose these obervables to be electric and magnetic fluxes. When one then tries to minimize the fluctuations of these obervables the parameters ϵ and a (the parameter that appears in $t = \ell_p^2/a^2$, we have chosen $a_e = const.$ for simplicity) get locked at $a \approx L$ and $\epsilon = \ell_p^{\alpha} L^{1-\alpha}$ for some $0 < \alpha < 1$ which in that case takes the value $\alpha = 1/6$. These considerations suggest the following conclusions:

1) *Three Scales*
   There are altogether three scales, the microscopic Planck scale ℓ_p, the mesoscopic scale ϵ and the macroscopic scale L. Since $\ell_p \ll L$ we have $\ell_p \ll \epsilon \ll L$ provided that (as in this case) α is not too close to the values $0, 1$.

2) *Geometric Mean*
   The mesoscopic scale takes a *geometric mean* between the microscopic and macroscopic scales. In particular, it lies well above the microscopic scale ℓ_p in contrast to the geometric weave states. The reason for this is that not only electric fluxes had to be well approximated but also magnetic ones: The weave states are basically spin-network functions which in turn are very similar to momentum eigenfunctions. Since then electric fluxes are very sharply peaked, magnetic ones are not peaked at all due to the Heisenberg uncertainty relation. This can best be seen by the observation that $< T_s, (\hat{h}_p)_{AB} T_s >= 0$ for any spin-network state and any $A, B = 1, 2$ (and therefore also $\omega_{q^0,\lambda,j,I}(\hat{h}_p) = 0$ for the statistical weave) which is an unacceptable expectation value since $\hat{h}_p$ should be $SU(2)-$valued. In order to approximate holonomies *one must take an average over large numbers of spins.* This is precisely what our coherent states do. As a consequence, the elementary observables, those that are defined at the

smallest scale which still allows semiclassical behaviour, are now defined at scales not smaller than $\epsilon \gg \ell_p$.

3) *Continuum Limit*
Notice that all our states and operators are defined in the continuum, therefore no continuum limit has to be taken. Yet, the scale ϵ could be associated with a measure for closeness to the continuum in which the graphs with which we probe operators tend to the continuum. The relation $\epsilon = \ell_p^\alpha L^{1-\alpha}$ reveals that not only one cannot take $\epsilon \to 0$ at finite ℓ_p because fluctuations would blow up, but also that the *"continuum limit" $\epsilon \to 0$ and the classical limit $\ell_p \to 0$ get synchronized.*

We expect many of those properties to hold generically for any semiclassical states that one may want to build for canonical quantum general relativity and that the extensive proofs in [208] will be useful for a whole class of states of this kind.

Remark:
Let us come back once more to the issue of using kinematical rather than dynamical coherent states. We already said that the full solution to the Hamiltonian constraint is not known at the moment and that even the operator with respect to which we want to compute these solutions is not under sufficient control. Therefore in a first step we must use kinematical states in order to make sure that we have the correct operator. Suppose then that we would have found the correct operator, then certainly the king's way of doing things would be to work with dynamical coherent states, but probably this would be highly impractical because the space of solutions for all constraints is very complicated (even classically we do not know all the solutions !). Thus, the poor man's way will be to consider kinematical coherent states ψ_m where m is a point on the constraint surface of the full phase space. The virtue of this is that the expectation value of full Dirac observables is *approximately gauge invariant* since

$$\delta_N < \psi_m, \hat{O}\psi_m >=< \psi_m, \frac{[\hat{H}(N), \hat{O}]}{i\hbar}\psi_m >\approx \{H(N), O\}(m) = 0$$

because O is a Dirac observable. Moreover

$$< \psi_m, \hat{O}\psi_m >\approx O(m) = O([m])$$

does not depend on the point m in the gauge orbit $[m]$ for the same reason. Thus, at least to zeroth order in $\hbar$ the expectation values of full Dirac observables and their infinitesimal dynamics should coincide whether we use kinematical or dynamical coherent states. This attitude is similar as in numerical classical gravity where one cannot just compute the time evolution of a given initial data set because for practical reasons one can only evolve approximately. The art is then to gain control on the error of these computations.

II.3.2.4 The Infinite Tensor Product Extension

Quantum field theory on curved spacetimes is best understood if the spacetime is actually flat Minkowski space on the manifold $M = \mathbb{R}^4$. Thus, when one wants to compute the low energy limit of canonical quantum general relativity to show that one gets the standard model (plus corrections) on a background metric one should do this first for the Minkowski background metric. Any classical metric is macroscopically non-degenerate. Since the quantum excitations of the gravitational field are concentrated on the edges of a graph, in order that, say, the expection values of the volume operator for any macroscopic region is non-vanishing and changes smoothly as we vary the region, the graph must fill the initial value data slice densely enough, the mean separation between vertices of the graph must be much smaller than the size of the region (everything is measured by the three metric, determined by the four metric to be approximated, in this case the Euclidean one). Now $\mathbb{R}^4$ is spatially non-compact and therefore such a graph must necessarily have an at least *countably infinite* number of edges whose union has *non-compact* range.

However, the Hilbert spaces in use for loop quantum gravity have as dense subspace the space of cylindrical functions labelled either by a piecewise analytic graph with a *finite* number of edges or by a so-called web, a piecewise smooth graph determined by the union of a *finite* number of smooth curves that intersect in a controlled way, albeit possibly a countably infinite number of times. Moreover, in both cases the edges or curves respectively are contained in *compact* subsets of the initial data hypersurface. These categories of graphs will be denoted by Γ_0^ω and Γ_0^∞ respectively where $\omega, \infty, 0$ stands for analytic, smooth and compactly supported respectively. Thus, the only way that the current Hilbert spaces can actually produce states depending on a countably infinite graph of non-compact range is by choosing elements in the closure of these spaces, that is, states that are countably infinite linear combinations of cylindrical functions.

The question is whether it is possible to produce semi-classical states of this form, that is, $\psi = \sum_n z_n \psi_{\gamma_n}$ where γ_n is either a finite piecewise analytic graph or a web, z_n is a complex number and we are summing over the intergers. It is easy to see that this is not the case : Minkowski space has the Poincaré group as its symmetry group and thus we will have to construct a state which is at least invariant under (discrete) spatial translations. This forces the γ_n to be translations of γ_0 and $z_n = z_0$. Moreover, the dependence of the state on each of the edges has to be the same and therefore the γ_n have to be mutually disjoint. It follows that the norm of the state is given by

$$||\psi||^2 = |z|^2([\sum_n 1][1 - |< 1, \psi_{\gamma_0} >|^2] + [\sum_n 1]^2 |< 1, \psi_{\gamma_0} >|^2)$$

where we assumed without loss of generality that $||\psi_{\gamma_0}|| = 1$ and we used the diffeomorphism invariance of the measure and 1 is the normalized constant state. By the Schwartz inequality the first term is non-negative and convergent only if $\psi_{\gamma_0} = 1$ while the second is non-negative and convergent only if $< 1, \psi_{\gamma_0} >= 0$. Thus the norm diverges unless $z = 0$.

This caveat is the source of its removal : We notice that the formal state $\psi := \prod_n \psi_{\gamma_n}$ really depends on an infinite graph and has unit norm if we formally compute it by $\lim_{N\to\infty} ||\prod_{n=-N}^N \psi_{\gamma_n}|| = 1$ using disjointness of the γ_n. The only problem is that this state is not any longer in our Hilbert space, it is not the Cauchy limit of any state in the Hilbert space : Defining $\psi_N := \prod_{n=-N}^N \psi_{\gamma_n}$ we find $| < \psi_N, \psi_M > | = | < 1, \psi_{\gamma_0} > |^{2|N-M|}$ so that ψ_N is not a Cauchy sequence unless $\psi_{\gamma_0} = 1$. However, it turns out that it belongs to the *Infinite Tensor Product (ITP) extension* of the Hilbert space.

To construct this much larger Hilbert space [209] we must first describe the class of graphs that we want to consider. We will consider graphs of the category Γ_σ^ω where σ now stands for countably infinite. More precisely, an element of Γ_σ^ω is the union of a countably infinite number of analytic, mutually disjoint (except possibly for their endpoints) curves called edges of compact or non-compact range which have no accumulation points of edges or vertices. In other words, the restriction of the graph to any compact subset of the hypersurface looks like an element of Γ_0^ω. These are precisely the kinds of graphs that one would consider in the thermodynamic limit of lattice gauge theories and are therefore best suited for our semi-classical considerations since it will be on such graphs that one can write actions, Hamiltonians and the like.

The construction of the ITP of Hilbert spaces is due to von Neumann [219] and already more than sixty years old. We will try to outline briefly some of the notions involved, see [209] for a concise summary of all definitions and theorems involved.

Let for the time being I be any index set whose cardinality $|I| = \aleph$ takes values in the set of non-standard numbers (Cantor's alephs). Suppose that for each $e \in I$ we have a Hilbert space $\mathcal{H}_e$ with scalar product $< .,. >_e$ and norm $||.||_e$. For complex numbers z_e we say that $\prod_{e\in I} z_e$ *converges* to the number z provided that for each positive number $\delta > 0$ there exists a finite set $I_0(\delta) \subset I$ such that for any other finite J with $I_0(\delta) \subset J \subset I$ it holds that $|\prod_{e\in J} z_e - z| < \delta$. We say that $\prod_{e\in I} z_e$ is *quasi-convergent* if $\prod_{e\in I} |z_e|$ converges. If $\prod_{e\in I} z_e$ is quasi-convergent but not convergent we define $\prod_{e\in I} z_e := 0$. Next we say that for $f_e \in \mathcal{H}_e$ the ITP $\otimes_f := \otimes_e f_e$ is a C_0 vector (and $f = (f_e)$ a C_0 sequence) if $|| \otimes_f || := \prod_{e\in I} ||f_e||_e$ converges to a non-vanishing number. Two C_0 sequences f, f' are said to be strongly resp. weakly equivalent provided that

$$\sum_e | < f_e, f'_e >_e -1| \text{ resp. } \sum_e || < f_e, f'_e >_e | - 1|$$

converges. The strong and weak equivalence class of f is denoted by $[f]$ and (f) respectively and the set of strong and weak equivalence classes by $\mathcal{S}$ and $\mathcal{W}$ respectively. We define the ITP Hilbert space $\mathcal{H}^\otimes := \otimes_e \mathcal{H}_e$ to be the closed linear span of all C_0 vectors. Likewise we define $\mathcal{H}^\otimes_{[f]}$ or $\mathcal{H}^\otimes_{(f)}$ to be the closed linear spans of only those C_0 vectors which lie in the same strong or weak equivalence class as f. The importance of these notions is that the determine much of the structutre of $\mathcal{H}^\otimes$, namely :
1) All the $\mathcal{H}^\otimes_{[f]}$ are isomorphic and mutually orthogonal.
2) Every $\mathcal{H}^\otimes_{(f)}$ is the closed direct sum of all the $\mathcal{H}^\otimes_{[f']}$ with $[f'] \in \mathcal{S} \cap (f)$.
3) The ITP $\mathcal{H}^\otimes$ is the closed direct sum of all the $\mathcal{H}^\otimes_{(f)}$ with $(f) \in \mathcal{W}$.
4) Every $\mathcal{H}^\otimes_{[f]}$ has an explicitly known orthonormal von Neumann basis.
5) If s, s' are two different strong equivalence classes in the same weak one then there exists a unitary operator on $\mathcal{H}^\otimes$ that maps $\mathcal{H}^\otimes_s$ to $\mathcal{H}^\otimes_{s'}$, otherwise such an operator does not exist, the two Hilbert spaces are unitarily inequivalent subspaces of $\mathcal{H}^\otimes$.
Notice that two isomorphic Hilbert spaces can always be mapped into each other such that scalar products are preserved (just map some orthonormal bases) but here the question is whether this map can be extended unitarily to all of $\mathcal{H}^\otimes$. Intuitively then, strong classes within the same weak classes describe the same physics, those in different weak classes describe different physics such as an infinite difference in energy, magnetization, volume etc. See [220] and references therein for illustrative examples.

Next, given a (bounded) operator a_e on $\mathcal{H}_e$ we can extend it in the natural way to $\mathcal{H}^\otimes$ by defining $\hat{a}_e$ densely on C_0 vectors through $\hat{a}_e \otimes_f = \otimes_{f'}$ with $f'_{e'} = f_{e'}$ for $e' \neq e$ and $f'_e = a_e f_e$. It turns out that the algebra of these extended operators for a given edge is automatically a von Neumann algebra [5, 39, 142, 143, 144] for $\mathcal{H}^\otimes$ (a weakly closed subalgebra of the algebra of bounded operators on a Hilbert space) and we will call the weak closure of all these algebras the von Neumann algebra $\mathcal{R}^\otimes$ of local operators. This way, adjointness relations and canonical commutation relations (Weyl algebra) are preserved.

Given these notions, the strong equivalence class Hilbert spaces can be characterized further as follows. First of all, for each $s \in \mathcal{S}$ one can find a representant $\Omega^s \in s$ such that $||\Omega^s|| = 1$. Moreover, one can show that $\mathcal{H}^\otimes_s$ is the closed linear span of those C_0 vectors $\otimes_{f'}$ such that $f'_e = \Omega^s_e$ for all but finitely many e. In other words, the strong equivalence class Hilbert spaces are irreducible subspaces for $\mathcal{R}^\otimes$, Ω^s is a cyclic vector for $\mathcal{H}^\otimes_s$ on which the local operators annihilate and create local excitations and thus, if I is countable, $\mathcal{H}^\otimes_s$ is actually separable. We see that we make naturally contact with Fock space structures, von Neumann algebras and their factor type classification [39] (modular theory) and algebraic quantum field theory [5]. The algebra of operators on the ITP which are not local do not have an immediate interpretation but it is challenging that they map between different weak equivalence classes and thus change the physics in a drastic way.

A number of warnings are in order :

1) Scalar multiplication is not multi-linear! That is, if f and $z\cdot f$ are C_0 sequences where $(z\cdot f)_e = z_e f_e$ for some complex numbers z_e then $\otimes_f = (\prod_e z_e)\otimes_f$ is in general wrong, it is true if and only if $\prod_e z_e$ converges.

2) Unrestricted use of the associative law of tensor products is false ! Let us subdivide the index set I into mutually disjoint index sets $I=\cup_\alpha I_\alpha$ where α runs over some other index set A. One can now form the different ITP $\mathcal{H}'^{\otimes} = \otimes_\alpha \mathcal{H}^{\otimes}_\alpha$, $\mathcal{H}^{\otimes}_\alpha = \otimes_{e\in I_\alpha}\mathcal{H}_e$. Unless the index set A is finite, a generic C_0 vector of $\mathcal{H}'^{\otimes}$ is orthogonal to all of $\mathcal{H}^{\otimes}$. This fact has implications for quantum gravity which we outline below.

Let us now come back to canonical quantum general relativity. In applying the above concepts we arrive at the following surprises :

i) First of all, we fix an element $\gamma\in\Gamma^\omega_\sigma$ and choose the countably infinite index set $E(\gamma)$, the edge set of γ. If $|E(\gamma)|$ is finite then the ITP Hilbert space $\mathcal{H}^{\otimes}_\gamma := \otimes_{e\in E(\gamma)}\mathcal{H}_e$ is naturally isomorphic with the subspace $\mathcal{H}^0_\gamma$ of $\mathcal{H}^0$ obtained as the closed linear span of cylinder functions over γ. However, if $|E(\gamma)|$ is truly infinite then a generic C_0 vector of $\mathcal{H}^{\otimes}_\gamma$ is orthogonal to any possible $\mathcal{H}^0_{\gamma'}$, $\gamma'\in\Gamma^\omega_0$. Thus, even if we fix only one $\gamma\in\Gamma^\omega_\sigma$, the total $\mathcal{H}^0$ is orthogonal to almost every element of $\mathcal{H}^{\otimes}_\gamma$.

ii) Does $\mathcal{H}^{\otimes}_\gamma$ have a measure theoretic interpretation as an L_2 space ? By the Kolmogorov theorem [58] the infinite product of probability measures is well defined and thus one is tempted to identify $\mathcal{H}^{\otimes}_\gamma = \otimes_e L_2(SU(2), d\mu_H)$ with $\mathcal{H}^{0\prime}_\gamma := L_2(\times_e SU(2), \otimes_e d\mu_H)$. However, this cannot be the case, the ITP Hilbert space is non-separable (as soon as $\dim(\mathcal{H}_e) > 1$ for almost all e and $|E(\gamma)|=\infty$) while the latter Hilbert space is separable, in fact, it is the subspace of $\mathcal{H}^0$ consisting of the closed linear span of cylindrical functions over γ' with $\gamma'\in\Gamma^\omega_0\cap E(\gamma)$.

iii) Yet, there is a relation between $\mathcal{H}^{\otimes}_\gamma$ and $\mathcal{H}^0$ through the inductive limit of Hilbert spaces : We can find a directed sequence of elements $\gamma_n\in\Gamma^\omega_0\cap E(\gamma)$, that is, $\gamma_m\subset\gamma_n$ for $m\le n$, such that γ is its limit in Γ^ω_σ. The subspaces $\mathcal{H}^0_{\gamma_n}\subset\mathcal{H}^0$ are isometric isomorphic with the subspaces of $\mathcal{H}^{\otimes}_\gamma$ given by the closed linear span of vectors of the form $\psi_{\gamma_n}\otimes[\otimes_{e\in E(\gamma-\gamma_n)}1]$ where $\psi_{\gamma_n}\in\mathcal{H}^0_{\gamma_n}\equiv\mathcal{H}^{\otimes}_{\gamma_n}$ which provides the necessary isometric monomorphism to display $\mathcal{H}^{\otimes}_\gamma$ as the inductive limit of the $\mathcal{H}^0_{\gamma_n}$.

vi) So far we have looked only at a specific $\gamma\in\Gamma^\omega_\sigma$. We now construct the total Hilbert space

$$\mathcal{H}^{\otimes} := \overline{\cup_{\gamma\in\Gamma^\omega_\sigma}\mathcal{H}^{\otimes}_\gamma}$$

equipped with the natural scalar product derived in [209]. This is to be compared with the Hilbert space

$$\mathcal{H}^0 := \overline{\cup_{\gamma\in\Gamma^\omega_0}\mathcal{H}^0_\gamma} = \overline{\cup_{\gamma\in\Gamma^\omega_\sigma}\mathcal{H}^{0\prime}_\gamma}$$

The identity in the last line enables us to specify the precise sense in which $\mathcal{H}^0\subset\mathcal{H}^{\otimes}$: For any $\gamma\in\Gamma^\omega_\sigma$ the space $\mathcal{H}^{0\prime}_\gamma$ is isometric isomorphic as specified in iii) with the strong equivalence class Hilbert subspace $\mathcal{H}^{\otimes}_{\gamma,[1]}$ where $1_e=1$ is the constant function equal to one. Thus, the Hilbert $\mathcal{H}^0$ space describes the local excitations of the "vacuum" Ω^0 with $\Omega^0_e=1$ for any possible analytic path e.

Notice that both Hilbert spaces are non-separable, but there are two sources of non-separability : the Hilbert space $\mathcal{H}^0$ is non-separable because Γ_0^ω has uncountable infinite cardinality. This is also true for the ITP Hilbert space but it has an additional character of non-separability : even for fixed γ the Hilbert space $\mathcal{H}_\gamma^\otimes$ splits into an uncountably infinite number of mutually orthogonal strong equivalence class Hilbert spaces and $\mathcal{H}_\gamma^{0\prime}$ is only one of them.

v) Recall that spin-network states [10] form a basis for $\mathcal{H}^0$. The result of iv) states that they are no longer a basis for the ITP. The spin-network basis is in fact the von Neumann basis for the strong equivalence class Hilbert space determined by $[\Omega^0]$ but for the others we need uncountably infinitely many other bases, even for fixed γ. The technical reason for this is that, as remarked above, the unrestricted associativity law fails on the ITP.

We would now like to justify this huge blow up of the original Hilbert space $\mathcal{H}^0$ from the point of view of physics. Clearly, there is a blow up only when the initial data hypersurface is non-compact as otherwise $\Gamma_0^\omega = \Gamma_\sigma^\omega$. Besides the fact that like $\mathcal{H}^0$ it is another solution to implementing the adjointness – and canonical commutation relations, we have the following:

a) Let us fix $\gamma \in \Gamma_\sigma^\omega$ in order to describe semi-classical physics on that graph as outlined above. Given a classical initial data set m we can construct a coherent state $\psi_{\gamma,m}$ which in fact is a C_0 vector $\otimes_{\psi_m}^\gamma$ for $\mathcal{H}_\gamma^\otimes$ of unit norm. This coherent state can be considered as a "vacuum" or "background state" for quantum field theory on the associated spacetime. As remarked above, the corresponding strong equivalence class Hilbert space $\mathcal{H}_{\gamma,[\psi_m]}^\otimes$ is obtained by acting on the "vacuum" by local operators, resulting in a space isomorphic with the familar Fock spaces and which is separable. In this sense, the fact that $\mathcal{H}_\gamma^\otimes$ is non-separable, being an uncountably infinite direct sum of strong equivalence class Hilbert spaces, could simply account for the fact that in quantum gravity *all vacua have to be considered simultaneously, there is no distinguished vauum as we otherwise would introduce a backgrond dependence into the theory.*

b) The Fock space structure of the strong equivalence classes immediately suggests to try to identify suitable excitations of $\psi_{\gamma,m}$ as graviton states propagating on a spacetime fluctuating around the classical background deteremined by m [211].
Also, it is easy to check whether for different solutions of Einstein's equations the associated strong equivalence classes lie in different weak classes and are thus physically different. For instance, preliminary investigations indicate that Schwarzschild black hole spacetimes with different masses lie in the *same* weak class. Thus, *unitary* black hole evaportation and formation seems not to be excluded from the outset.

c) From the point of view of $\mathcal{H}_\gamma^{0\prime}$ the Minkowski coherent state is an everywhere excited state like a thermal state, the strong classes $[\Omega^0]$ and $[\psi_m]$ for Minkowski data m are orthogonal and lie in different weak classes. The state Ω^0 has no obvious semi-classical interpretation in terms of coherent states for any classical spacetime.

d) It is easy to see that the GNS Hilbert space used in [202] is isometric isomorphic with a strong equivalence class Hilbert space of our ITP construction. Thus, our ITP framework collects a huge class of representations in the "folium" [5] of the Hilbert space $\mathcal{H}^0$ and embeds them isometrically into one huge Hilbert space $\mathcal{H}^\otimes$, thus we have now *an inner product between different GNS Hilbert spaces*! This demonstrates the power of this framework because inner products between different GNS Hilbert spaces are normally not easy to motivate.

II.3.3 Photon Fock States on $\overline{\mathcal{A}}$

In [212] Varadarajan investigated the question in which sense the the techniques of $\overline{\mathcal{A}}, \mu_0$, which in principle apply to any gauge field theory of connections for compact gauge groups, can be used to describe the Fock states of Maxwell theory. This is not at all an academic question because presumably one wants to couple Maxwell theory to gravity also in such a background independent representation as, in fact we have indicated in section II.2. Moreover, linearized gravity can be described in terms of connections as well [213] where it becomes effectively a $U(1)^3$ Abelean gauge theory just like Maxwell theory. Both theories are, of course, ordinary free field theories on a Minkowski background.

Varadarajan succeeded in displaying Fock states within the framework of $\overline{\mathcal{A}}, \mu_0$ in a very precise way. The crucial observation, unfortunately only valid if the gauge group is Abelean, is the following isomorphism between two different Poisson subalgebras of the Poisson algebra on $\mathcal{M}$: Consider a one-parameter family of test functions of rapid decrease which are regularizations of the $\delta-$distribution, for instance

$$f_r(x,y) = \frac{e^{-\frac{||x-y||^2}{2r^2}}}{(\sqrt{2\pi}r)^3} \tag{II.3.3. 1}$$

where we have made use of the Euclidean spatial background metric. Given a path $p \in \mathcal{P}$ we denote its form factor by

$$X^a_p(x) := \int_0^1 dt \dot{p}^a(t)\delta(x,p(t)) \tag{II.3.3. 2}$$

The smeared form factor is defined by

$$X^a_{p,r}(x) := \int d^3y f_r(x,y) X^a_p(y) = \int_0^1 dt \dot{p}^a(t) f_r(x,p(t)) \tag{II.3.3. 3}$$

which is evidently a test function of rapid decrease. Notice that a $U(1)$ holonomy maybe written as

$$h_p(A) := e^{i\int d^3x X^a_p(x) A_a(x)} \tag{II.3.3. 4}$$

and we can define a smeared holonomy by

$$h_{p,r}(A) := e^{i\int d^3x X^a_{p,r}(x) A_a(x)} \tag{II.3.3. 5}$$

Likewise we may define smeared electric fields as

$$E^a_r(x) := \int d^3y f_r(x,y) E^a(y) \tag{II.3.3. 6}$$

If we denote by q the electric charge (notice that in our notation $\alpha = \hbar q^2$ is the fine structure constant), then we obtain the following Poisson subalgebras: On the one hand we have smeared holonomies but unsmeared electric fields with

$$\{h_{p,r}, h_{p',r}\} = \{E^a(x), E^b(y\} = 0, \ \ \{E^a(x), h_{p,r}\} = iq^2 X^a_{p,r}(x) h_{p,r} \tag{II.3.3. 7}$$

and on the other hand we have unsmeared holonomies but smeared electric fields with

$$\{h_p, h_{p'}\} = \{E^a_r(x), E^b_r(y\} = 0, \ \ \{E^a_r(x), h_p\} = iq^2 X^a_{p,r}(x) h_p \tag{II.3.3. 8}$$

Thus the two Poisson algebras are ismorphic and also the * relations are isomorphic, both $E^a(x), E^a_r(x)$ are real valued while both $h_p, h_{P,r}$ are $U(1)$ valued. Thus, as abstract $^*-$ Poisson algebras these two algebras are indistinguishable and we may ask if we can find different representations of it. Even better, notice that $h_{p,r}h_{p',r} = h_{p\circ p',r}$, $h_{p,r}^{-1} = h_{p^{-1},r}$ so the smeared holonomy algebra is also isomorphic to the unsmeared one. It is crucial to point out that the right hand side of both (II.3.3. 7), (II.3.3. 8) is a cylindrical function again only in the Abelean case, see section I.3.1. Therefore all that follows *is not true for* $SU(2)$.

Now we know that the unsmeared holonmy algebra is well represented on the Hilbert space $\mathcal{H}^0 = L_2(\overline{\mathcal{A}}, d\mu_0)$ while the smeared holonomy algebra is well represented on the Fock Hilbert space $\mathcal{H}_F = L_2(\mathcal{S}', d\mu_F)$ where $\mathcal{S}'$ denotes the space of divergence free, tempered distributions and μ_F is the Maxwell-Fock measure. These measures are completely characterized by their generating functional

$$\omega_F(\hat{h}_{p,r}) := \mu_F(h_{p,r}) = e^{-\frac{1}{4\alpha}\int d^3x X^a_{p,r}(x)\sqrt{-\Delta}^{-1} X^b_{p,r}\delta_{ab}} \tag{II.3.3. 9}$$

since finite linear combinations of the $h_{p,r}$ are dense in $\mathcal{H}_F$ [212]. Here $\Delta = \delta^{ab} = \partial_a\partial_b$ denotes the Laplacian. Here we have taken a loop p rather than an open path so that $X_{p,r}$ is transversal. Also unsmeared electric fields are represented through the Fock state ω_F by

$$\omega_F(\hat{h}_{p,r}\hat{E}^a(x)\hat{h}_{p',r}) = -\frac{\alpha}{2}[X^a_{p,r}(x) - X^a_{p',r}(x)]\omega_F(\hat{h}_{p\circ p',r}) \tag{II.3.3. 10}$$

and any other expectation value follows from these and the commutation relations.

Since ω_F defines a positive linear functional we may define a new representation of the algebra h_p, E^a_r by

$$\omega_r(\hat{h}_p) := \omega_F(\hat{h}_{p,r}) \text{ and } \omega_r(\hat{h}_p\hat{E}^a_r(x)\hat{h}_{p'}) := \omega_F(\hat{h}_{p,r}\hat{E}^a(x)\hat{h}_{p',r}) \tag{II.3.3. 11}$$

called the $r-$Fock representation. In order to see whether there exists is a measure μ_r on $\overline{\mathcal{A}}$ that represents ω_r in the sense of the Riesz representation theorem we must check that ω_r is a positive linear functional on $C(\overline{\mathcal{A}})$. This can be done [212]. In [214] Velhinho has computed explicitly the cylindrical projections of this measure and showed that the one – parameter family of measures μ_r are expectedly mutually singular with respect to each other and with respect to the uniform measure μ_0. Thus, none of these Hilbert spaces is contained in any other. In fact, we have a natural map

$$\Theta_r : \mathcal{S}' \to \overline{\mathcal{A}/\mathcal{G}};\ A \mapsto \Theta_r(A) \text{ where } [\Theta_r(A)](p) := e^{i\int d^3x X^a_{p,r}A_a(x)} \tag{II.3.3. 12}$$

and Velhinho showed that $\mu_r = (\Theta_r)_*\mu_F$ is just the push-forward of the Fock measure.

Recall that the Fock vacuum Ω_F is defined to be the zero eigenvalue coherent state, that is, it is annihilated by the annihilation operators

$$\hat{a}(f) := \frac{1}{\sqrt{2\alpha}}\int d^3x f^a[\sqrt[4]{-\Delta}\hat{A}_a - i(\sqrt[4]{-\Delta})^{-1}\hat{E}^a] \tag{II.3.3. 13}$$

where f^a is any transversal smearing field. We then have in fact that $\omega_F(.) = <\Omega_F, .\Omega_F>_{\mathcal{H}_F}$, that is Ω_F is the cyclic vector that is determined by ω_F through the GNS construction. The idea is now the following: From (II.3.3. 11) we see that we can easily answer any question in the $r-$Fock representation which has a preimage in the Fock representation, we just have to replace everywhere $h_{p,r}, E^a(x)$ by $h_p, E^a_r(x)$. Since in the $r-$Fock representations only exponentials of connections are defined, we should exponentiate the annihilation operators and select the Fock vacuum through the condition

$$e^{i\hat{a}(f)}\Omega_F = \Omega_F \tag{II.3.3. 14}$$

In particular, choosing $f = \sqrt{2\alpha}(\sqrt[4]{-\Delta})^{-1}X_{p,r}$ for some loop p we get

$$e^{\int d^3x X^a_{p,r}[i\hat{A}_a + (\sqrt{-\Delta})^{-1}\hat{E}^a]}\Omega_F = \Omega_F \tag{II.3.3. 15}$$

Using the commutation relations and the Baker – Campell – Hausdorff formula one can write (II.3.3. 15) in terms of $\hat{h}_{p,r}$ and the exponential of the electric field appearing in (II.3.3. 15) times a numerical factor. The resulting expression can then be translated into the $r-$Fock representation.

This was Varadarajan's idea. He found that in fact there is no state in $\mathcal{H}^0$ which satisfies the translated analogue of (II.3.3. 15) but that there exists a distribution that does (we must translate (II.3.3. 15) first into the dual action to compute that distribution). It is given (up to a constant) by

$$\Omega_r = \sum_s e^{-\frac{\alpha}{2}\sum_{e,e'\in E(\gamma(s))} G^r_{e,e'} n_e(s) n_{e'}(s)} T_s < T_s, . >_{\mathcal{H}^0} \tag{II.3.3. 16}$$

where $s = (\gamma(s), \{n_e(s)\}_{e\in E(\gamma(s))})$ denotes a charge network (the $U(1)$ analogue of a spin network) and

$$G^r_{e,e'} = \int d^3x X^a_{e,r} \sqrt{-\Delta}^{-1} X^b_{e',r} \delta^T_{ab} \tag{II.3.3. 17}$$

where $\delta^T_{ab} = \delta_{ab} - \partial_a \Delta^{-1}\partial_b$ denotes the transverse projector.

Several remarks are in order concerning this result:

1) *Distributional Fock States*
$n-$particle state excitations of the state Ω_F (and also coherent states [215]) can be easily translated into distributional $n-$particle states (coherent states) by using Varadarjan's prescription above. Thus, we get in fact a **Varadarajan map**

$$V:\ (\mathcal{H}_F, \mathcal{L}(\mathcal{H}_F) \mapsto (\mathcal{D}^*, \mathcal{L}'(\mathcal{D})) \tag{II.3.3. 18}$$

Of course, none of the image states is normalizable with respect to μ_0 and this raises the question in which sense the kinematical Hilbert space is useful at all in order to do semi-classical analysis. One can *in this case define* a new scalar product on these distributions simply by

$$< V\cdot\psi, V\cdot\psi' >_r := < \psi, \psi' >_F \tag{II.3.3. 19}$$

In particular we obtain $< \Omega_r, . \Omega_r >_r = \omega_r$ so Ω_r can be interpreted as the GNS cyclic vector underlying ω_r. With respect to this inner product one can now perform semi-classical analysis. Of course, in the non-Abelean case a Varadarjan map is not available at this point.

2) *Electric Flux Operators*
In the non-Abelean theory it was crucial not to work with electrical fields smeared in D dimensions but rather with those smeared in $D-1$ dimensions. However, $(D-1)-$smeared electrical fields have no pre-image under V and in fact Velhinho showed that there is no electric flux operator in the $r-$Fock representation as to be expected. This seems to be an obstruction to transfer the Varadarajan map to the non – Abelean case.

3) *Comparison with Heat Kernel Coherent States*
Formulas (II.3.2. 20) and (II.3.3. 16) look very similar to each other (see also [215]). We can write (II.3.3. 16) more suggestively as

$$\Omega_r = \sum_s e^{\frac{\alpha}{2}\sum_{e,e'\in E(\gamma(s))} G^r_{e,e'} R_e R_{e'}} T_s < T_s, . >_{\mathcal{H}^0} \tag{II.3.3. 20}$$

where R_e are right invariant vector fields on $U(1)$. This formula just asks to be analytically continued in order to arrive at a coherent state because it looks like (II.3.2. 20). The deeper origin of this apparent coincidence will be unravelled in [218] where it will be shown that the Varadarajan coherent distributions are a special case of the general formula (II.3.2. 21).

In [215] it is speculated that one should generalize (II.3.3. 20) in the obvious way to the non-Abelean case by replacing charge nets by spin nets and $R_e R_{e'}$ by $R^j_e R^j_{e'}$ and to use the associated cut-off states (called "shadows" there) for semi-classical analysis. However, it is unclear whether these shadows have similarly nice properties as the cut-off states introduced in [207, 208, 209, 210] because the metric $G^r_{ee'}$ is not diagonal. Also it is unclear how one should then define non-Abelean Fock states. Finally it is not clear what the interpretation of the complexified group label should be without which a semiclassical ineterpretation of those states is out of reach.

4) *Other Operators*
One should not forget that important operators of Maxwell theory such as the Hamiltonian operator are expressed as polynomials of *un – exponentiated* annihilation and creation operators. However, such operators are not defined neither in the $r-$Fock representation nor in $\mathcal{H}^0$. In [211] we will show how to circumvent that problem.

II.3.4 Applications

Beyond merely checking whether we have a quantum theory of the correct classical theory, namely general relativity coupled to all known matter, quantum gravity has certainly a huge impact on the whole structure of physics. For instance, if the picture drawn in section II.2 is correct, then one must do quantum field theory on one-dimensional polymer like structures rather than in a higher dimensional manifold, presumably the ultraviolet divergences disappear and while there are still bare and renormalized charges, masses etc. the bare charges will presumably be finite while the renormalized charges should better be called effective charges because they simply take into account physical screening effects.

Quantum gravity effects are notoriously difficult to measure because the Planck length is so incredibly tiny. It may therefore come as a surprise that recently physicists have started to seriously discuss the possibility to measure quantum gravity effects, mostly from astrophysical data and gravitational wave detectors [221]. See also the discussion in the extremely beautiful review by Carlip [222] and references therein. The challenge is to compute these effects within quantum general relativity. First pioneering steps towards the computation of the so-called $\gamma-$*ray burst effect* have been made, to date mostly at a phenomenological level, in [223] for photons and [224] for neutrinos. A more detailed analysis based on the coherent states proposed in [207, 208] will appear in [211].

This is not the place to give a full-fledged account of these developments, so we will restrict ourselves to presenting the main ideas for the $\gamma-$ray burst effect.

A $\gamma-$ray burst is a light signal of extremely high energetic photons (up to 1 TeV !) that travelled over cosmological distances (say 10^9 years). What is interesting about them is that the signal is like a flash, that is, the intensity decays on the order of 10^{-3}s. The astrophysical origin of these bursts is still under debate (see the references in [224]) and we will have nothing to add on this debate here. What is important though is that these photons probe the discrete (polymer) structure of spacetime the more, the more energy they have which should lead to an energy dependent velocity of light (dispersion) very similar to the propagtion of light in cristals. More specifically, if one plots the time signal of events as measured by a atmospherical Cerenkov light detector [225] within two disjoint energy channels $[E_1 - \Delta E, E_1 + \Delta E]$ and $[E_2 - \Delta E, E_2 + \Delta E]$ then one expects a time difference in the peak of these signals given by

$t_2 - t_1 = \xi \frac{L}{c(0)}[(E_2/E_p)^\alpha - (E_1/E_p)^\alpha]$ where L is the difference from the source (measured by the red shift of the galaxy), $c(0)$ is the vacuum speed of light, E_p is the effective Planck scale energy of the order of m_p and α, ξ are theory dependent constants of the order unity. If $\alpha = \xi = 1$, $E_p = m_p$ and $E_2 - E_1$ =1 TeV then for $L = 10^9$ lightyears we get travel time differences of the order of 10^2s which is much larger than the duration of the peak. At present, the sensitivity of available detectors is way below such a resoltion of ms mainly because no detectors ahve been built for this specific purpose but the construction of better detectors is on the way [224].

One may object that 1) quantum field theory effects from other interactions should be much stronger than quantum gravity effects so that this effect would not test so much quatum gravity but rather quantum field theory on Minkowski space, 2) there are many possible astrophysical disturbances that can cause dispersion such as interstellar dust and 3) it is not clear that the photons of different energies have been emitted simultaneously.

The answer is as follows:

1) is ecluded by definition of quantum field on Minkowski space: Such a theory is Poincaré invariant by construction while an energy dependent dispersion breaks Lorentz invariance. We see that the effect is *non-perturbative* because in any perturbative approach to quantum gravity one treats gravity like the other inetactions as a quantum field theory on a Minkowski background.

2) is excluded by the fact that the effect gets stronger with higher energy while diffraction at dust gets weaker: The scale of dust or gas molecules is transparent for such highly energetic photons.

3) is apparently excluded by model computations in astrophysics [225] for the known scenarios that lead to the $\gamma-$ray burst effect.

How would one then compute the effect within quantum general relativity ? Basically, one would look at quantum Einstein-Maxwell theory and consider states of the form $\psi_E \otimes \psi_M$ where ψ_E is a fixed coherent state for the gravitational degrees of freedom, peaked at Minkowski initial data and ψ_M is a quantum state for the Maxwell-field. Given the Einstein-Maxwell Hamiltonian

$$H_{EM} = \frac{1}{2e^2} \int d^3x \frac{q_{ab}}{\sqrt{\det(q)}} [E^a E^b + B^a B^b]$$

one would quantize it as described in section II.2 and then define an effective Maxwell Hamiltonian by

$$< \psi_M, \hat{H}_M^{eff} \psi'_M >_{\mathcal{H}_M} := < \psi_E \otimes \psi_M, \hat{H}_{EM} \psi_E \otimes \psi_M >_{\mathcal{H}_E \otimes \mathcal{H}_M}$$

At the moment we can do this computation only at the kinematical level but as outlined in section II.3.2 this should approximate the full dynamical computation and at least gives an idea for the size of the effect.

Whatever technique is finally being used to carry out this computation the mere existence of the effect is a *prediction* of any background independent approach to quantum gravity. In fact, the technical reason for existence of the effect is a corollary from the Heisenberg uncertainty relation: The quantum metric operators form a non-commuting set of operators (they depend both on magnetic and electric degrees of freedom) so that it is not possible to diagonalize them simultaneously. The best one can do is to construct an approximate eigenstate for all of them (namely a coherent state) but that state can then not be exactly Poincaré invariant, only approximately.

There are countless other applications of semiclassical states such as an approach to quantum black holes from first principles and a corresponding computation of the Hawking effect that takes full account of the backreaction of the gravitational field towards infalling matter which at the horizon becomes infinitely blue shifted so that quantum gravity effects are no longer neglible.

II.4 Further Research Directions

In the second last section of this review we will describe briefly three more major research directions within Canonical Quantum General Relativity: Spin Foam Models, Quantum Black Holes and Interfaces between Canonical Quantum General Relativity and String Theory. To be sure, all three topics deserve to be treated in a chapter of their own, however, our presentation will be short since a thorough treatment would require three additional reviews in their own right plus extra background material in additional appendices which would explode the already huge length of this review. Luckily, nice, pretty self-contained, review articles, at least for the two first programmes, already exist:

For an introduction to spin-foam models we recommend the really beautiful article by Baez [226] which contains an almost complete and up to date guide to the literature and the historical development of the subject. See also the article by Barrett [227] for the closely related subject of state sum models. A summary of the classical and quantum aspects of so-called *isolated horizons*, a local generalization of event horizons that is used in black hole entropy calculations within quantum general relativity, can be found in [228]. The pivotal papers that describe the details of the classical and quantum formulation respectively are [229] and [230] respectively.

II.4.1 Spin Foam Models

The prototype of spinfoam models are state sum models that had been extensivley studied [231] within the context of topological quantum field theories [232] long before spin foam models arose within quantum gravity. The concrete connection of state sum models with canonical quantum gravity was made by Reisenberger and Rovelli in their seminal paper [233] where they used the (Euclidean version of the) Hamiltonian constraint described in section II.1 in order to write down a path integral formulation of the the theory. Roughly speaking, this works as follows:
A heuristic method of how to solve the Hamiltonian constraint is to take any kinematical state ψ and to map it to $\delta(\hat{H})\psi$ where $\delta(\hat{H})=\prod_{x\in\sigma}\delta(\hat{H}(x)$. This is of course quite formal since neither the $\hat{H}(x)$ are self-adjoint nor mutually commuting. It is anyway a formal solution to the Hamiltonian constraint if we treat the Diffeomorphism constraint similarly because the algebra of deffeomorphisms and Hamiltonians is formally closed. Proceeding formally, we may define a path integral formulation of the $\delta-$distribution. Neglecting an (infinite) constant as usual we obtain the functional integral

$$\delta(\hat{H})=\int[dN]e^{i\int_\sigma d^3x N(x)\hat{H}(x)} \tag{II.4.1. 1}$$

This looks like a group averaging operation and we may try to define a physical inner product between physical states $\psi_{phys}:=\delta(\hat{H})\psi$ as

$$<\psi_{phys},\psi'_{phys}>_{phys}:=<\psi,\delta(\hat{H})\psi'>=\int_{\mathcal{N}}[dN]<\psi,e^{i\int_\sigma d^3x N(x)\hat{H}(x)}\psi'> \tag{II.4.1. 2}$$

where $\mathcal{N}$ is the set of all lapse functions on σ. In order to get time dependent lapse functions $\bar{N}(x,t)$ consider the set of lapse functions $\overline{\mathcal{N}}_N$ on M with $\int_{-T}^T dt\bar{N}(x,t)=N(x)$ for some $T>0$. Let also $\overline{\mathcal{N}}$ be the set of lapse functions over M. Then

$$\begin{aligned}
&\int_{\overline{\mathcal{N}}}[d\bar{N}]<\psi,e^{i\int_M d^4x N(x,t)\hat{H}(x)}\psi'> \qquad \text{(II.4.1. 3)}\\
=&\lim_{T\to\infty}\int_{\overline{\mathcal{N}}}[d\bar{N}]<\psi,e^{i\int_{-T}^T dt\int_\sigma d^3x N(x,t)\hat{H}(x)}\psi'>\\
=&\lim_{T\to\infty}\int_{\mathcal{N}}[dN]<\psi,e^{i\int_\sigma d^3x N(x)\hat{H}(x)}\psi'>[\int_{\overline{\mathcal{N}}}[d\bar{N}]\delta(\int_{-T}^T dt\bar{N}(x,t),N(x))]
\end{aligned}$$

Consider the integral

$$I_N^T := \int_{\bar{\mathcal{N}}} [d\bar{N}] \delta(\int_{-T}^{T} dt \bar{N}(x,t) = N(x)) \qquad \text{(II.4.1. 4)}$$

appearing in the square bracket in the last line of (II.4.1. 3). We claim that it is actually independent of $N(x)$. This can be verified by introducing the constant shift $\bar{N}(x,t) \mapsto \bar{N}(x,t) + \frac{N'(x)-N(x)}{2T}$ so that $I_N^T = I_{N'}^T = const.$. We conclude that (II.4.1. 3) and (II.4.1. 2) are proportional to each other (by an infinite constant $\lim_{T\to\infty} I_N^T$). The formula (II.4.1. 3) is then the starting point for formulating a path integral through the usual skeletonization process.

In any case we can now formally expand the exponent in (II.4.1. 2) and arrive at the following picture: Given two spin-network functions $T_s, T_{s'}$ we have

$$< T_{s,phys}, \psi'_{s',phys} >_{phys} := \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_{\mathcal{N}} [dN] < T_s, \hat{H}(N)^n T_{s'} > \qquad \text{(II.4.1. 5)}$$

Since $\hat{H}(N)$ is closed and densely defined on spin-network functions, the matrix elements of powers of the Hamiltonian constraint can be computed and since we integrate over all possible lapse functions the result is manifestly spatially diffeomorphism invariant. Of course, the result is badly divergent, but cutting off the integral over N somehow the following picture emerges: The power of $\hat{H}(N)^n$ corresponds to a discrete n time step evolution of an intial spin-net s' to a final one s. At each step $\hat{H}(N)$ changes the graph of the spin net s' according to the rules of section II.1. Let us associate a hypersurface with each time step and let the respective spin nets be embedded inside them. Connect the vertices of the spin-nets in subsequent hypersurfaces by dotted lines. Since $\hat{H}(N)$ adds edges to a graph, one of these dotted lines branches up at some intermediate point into two additional dotted lines which connect with the two newly created vertices.

We thus see that the quantum time evolution of edges become two-surfaces (bounded by one or two edges and two dotted lines), that is, *a spin foam.* Such kind of transition amplitudes are exactly of the form as considered earlier by Reisenberger already [234].

Thus, the canonical theory seems to suggest a bubble evolution not unlike the worldsheet formulation of string theory, although spin foams define a background independent string theory in which the worldsheet is not a smooth two-dimensional manifold but has necessarily (conical) singularities due to the fact that the Hamiltonian constraint acts non-trivially only at vertices in each time step.

In order to give mathematical meaning to these amplitudes one obviously has to look for a better definition of the path integral. One will therefore begin with stripping off all the particulars of the specific theory that describes quantum gravity and consider very general spin foam models and search for criteria when they converge and when they do not. Then, in a second step, one has to select among the converging ones the theory which describes quantum gravity (if any).

It turns out that a systematic starting point are the so-called BF topological field theories [232]. In $D+1$ dimensions these are described by an action ($D \geq 2$)

$$S_{BF} = \int_M \mathrm{Tr}(B \wedge F) \qquad \text{(II.4.1. 6)}$$

where B is a Lie(G) valued $(D-1)-$form in a vector bundle associated to a principal G bundle P under the adjoint representation and F is the curvature of a connection A over P. The trace operation is with respect to the the non-degenerate Cartan-Killing metric on Lie(G) (assuming G to be semi-simple), that is, basically the Kronecker symbol (up to normalization). The equations of motion are given by $F = DB = 0$ where D is the covariant differential determined by A (see section III.2). Thus A is constrained to be flat. The action has a huge symmetry, namely it is gauge invariant and invariant under $A \mapsto A,\ B \mapsto B + Df$ for any $(D-2)-$form f. Counting physical degrees of freedom it is easy to see that almost nothing is left, the theory has only a finite number of degrees of freedom, it is topological.

The connection with gravity is made through the Plebanski (first order) action (in this section we set $\kappa = 1$)

$$S_P = \int_M \mathrm{Tr}((*[e \wedge e]) \wedge F) \tag{II.4.1. 7}$$

Here $e = (e^j_\mu)$ denotes the co-(D+1)-bein and $*$ denotes the Hodge dual with respect to the internal metric η_{ij} which is just the Minkowski (Euclidean) metric for Lorentzian (Euclidean) general relativity with gauge group $SO(D,1)$ $(SO(D+1))$. More specifically

$$(*[e \wedge e])_{ij} := \frac{1}{(D-1)!}\epsilon_{ijk_1..k_{D-1}} e^{k_1} \wedge .. \wedge e^{k_{D-1}} \tag{II.4.1. 8}$$

and plugging this into (II.4.1. 7) one easily sees that (II.4.1. 7) equals the Einstein-Hilbert action for orientable M when A is the spin-connection of e (which is one of the equations of motion that one derives from (II.4.1. 7)). Thus we see that gravity is a BF theory modulo the constraint that B is in this case not an arbitrary $(D-1)-$form but rather has to satisfy the so-called *simplicity* constraint

$$B = *[e \wedge e] \tag{II.4.1. 9}$$

The idea for writing a path integral for general relativity is then the following: A lot is known about the path integral quantization of BF theory in three and four dimensions [231]. Thus, it seems to be advisable to consider general relativity as a BF theory in which the sum over histories is constrained by (II.4.1. 9). One might wonder how it can happen that a TQFT like BF theory with only a finite number of degrees of freedom plus additional constraints can give rise to a field theory like general relativity with an infinite number of degrees of freedom. The answer is that (II.4.1. 9) breaks a lot of the gauge invariance of BF theory so that gauge degrees of freedom become physical degrees of freedom. In order to sum over histories of B's and A's with the constraint (II.4.1. 9) we must first write it in a form in which only B's appear. The algebraic condition on B such that there exists e with (II.4.1. 9) satisfied has been systematically analyzed by Freidel, Krasnov and Puzio in [235]. It can be written for $D \geq 3$ as

$$\epsilon^{ijklm_1..m_{D-3}} B^{\mu\nu}_{ij} B^{\rho\sigma}_{kl} = \epsilon^{\mu\nu\rho\sigma\lambda_1..\lambda_{D-3}} c^{m_1..m_{D-3}}_{\lambda_1..\lambda_{D-3}} \tag{II.4.1. 10}$$

where c is any totally skew (in both sets of indices) tensor density and

$$B^{\mu\nu}_{ij} = \frac{1}{(D-1)!}\epsilon^{\mu\nu\rho_1..\rho_{D-1}} \eta_{ik}\eta_{jl} B^{kl}_{\rho_1..\rho_{D-1}} \tag{II.4.1. 11}$$

Actually for $D = 3$ there is another solution to (II.4.1. 10) besides (II.4.1. 9) given by

$$B = \pm e \wedge e \tag{II.4.1. 12}$$

but this solution gives rise again to a topological theory. The constraint (II.4.1. 10) is enforced by adding to the BF action a term of the form

$$\frac{1}{2}\int_M d^{D+1}x \Phi^{ijkl}_{\mu\nu\rho\sigma} B^{\mu\nu}_{ij} B^{\rho\sigma}_{kl} =: \frac{1}{2}\int_M \mathrm{tr}(B \wedge \Phi(B)) =: \int_M \Phi \cdot C \tag{II.4.1. 13}$$

where the Lagrange multiplier Φ is totally skew in both index sets and we have denoted the simplicity constraint by C.

Now the partition function for BF theory is given by

$$Z_{BF} = \int [dA\ dB] e^{i\int_M \mathrm{tr}(B\wedge F)} \propto \int [dA] \delta(F) \qquad \text{(II.4.1. 14)}$$

where for *either signature* the factor of i in front of the action has to be there in order to enforce the flatness constraint $\delta(F)$. That this defines the correct path integral (up to proper regularization) has been verified by independent methods, see [231, 232] and references therein. Since, from the point of view of BF theory, general relativity is a "perturbation" (with the role of the "free" theory being played by BF theory) with interaction term (II.4.1. 13) the partition function for general relativity should be given by

$$Z_P = \int [dA\ dB\ d\Phi] e^{i\int_M \mathrm{tr}(B\wedge [F+\frac{1}{2}\Phi(B)])} \propto \int [dA\ dB] \delta(C) e^{i\int_M \mathrm{tr}(B\wedge F)} \qquad \text{(II.4.1. 15)}$$

where the additional integral over the Lagrange multiplier enforces the simplicity constraint. Path integrals of the type (II.4.1. 15) were studied by Freidel and Krasnov [236] in terms of a generating functional

$$Z[J] := \int [dA\ dB] e^{i\int_M \mathrm{tr}(B\wedge [F+J])} \qquad \text{(II.4.1. 16)}$$

where J is a two-form current. It is easy to see that formally by a trick familiar from ordinary quantum field theory

$$Z_P = \int [d\Phi] \{ e^{i\frac{1}{2}\int_M \mathrm{tr}(\frac{\delta}{i\delta J}\Phi(\frac{\delta}{i\delta J})])} Z[J] \}_{J=0} \qquad \text{(II.4.1. 17)}$$

which could then be the starting point for perturbative expansions. Unfortunately, a truly systematic derivation of spin foam models for general relativity starting directly from (II.4.1. 17) is still missing.

We see that in order to define the partition function for general relativity we must first define the one for BF theory. Let us first consider the case that G is compact (Euclidean signature). Then the $\delta-$distribution $\delta(F)$ in (II.4.1. 14) can be interpreted as the condition that the holonomy of every contractible loop is trivial. Furthermore, in order to regularize the functional integral, we triangulate M, using some triangulation T and interpret the measure $[dA]$ as the uniform measure on $\overline{\mathcal{A}}$ restricted to T. Then the condition $F=0$ amounts to saying that $h_\alpha = 1_G$ where α is any contractible loop within T. Let $\pi_1'(T)$ be the generators of the contractible subgroup of the fundamental group of T. Hence the regulated BF partition function becomes

$$Z_{BF}(T) = \int_{\overline{\mathcal{A}}_T} d\mu_{0T}(A) \prod_{\alpha\in\pi_1'(T)} \delta(A(\alpha), 1_G) \qquad \text{(II.4.1. 18)}$$

and we can use the Peter&Weyl theorem in order to write the $\delta-$distribution as

$$\delta(h, 1_G) = \sum_{\pi\in\Pi} d_\pi \chi_\pi(h) \qquad \text{(II.4.1. 19)}$$

Now magically the integral (II.4.1. 18) is independent of the choice of triangulation which can be traced back to the fact that BF is a topological theory. The theory defined by (II.4.1. 18) is known as the Turarev-Viro state sum model for $D=2, G=SU(2)$ and as the Turarev-Ooguri-Crane-Yetter model in $D=3, G=SO(4)$. Actually (II.4.1. 18) is still divergent when one expands out the products of $\delta-$distributions but this can be taken care of by using a quantum group regularization at a root of unity which cuts off the sum over representations at those of bounded dimension.

Let us now turn to Euclidean gravity for $D = 3$. We somehow must invoke the simplicity constraint into (II.4.1. 18). The idea is to look at a canonical quantization of BF theory with the additional simplicity constraint imposed. This analysis has been started by by Barbieri [237] leading to the consideration of *quantum tetrahedra* and was completed by Baez and Barrett [238]. The result is as follows: Recall that $SO(4)$ is homomorphic with $SU(2) \times SU(2)$, therefore its irreducible representations can be labelled by two spin quantum numbers (j, j') ("left handed and right handed"). The simplicity constraint now amounts to the constraint $j = j'$ explaining the word "simplicity". This motivates to define the partition function for general relativity by restricting the sum in

$$\delta(h, 1_{SO(4)}) = \sum_{j,j'} d_{\pi_{j,j'}} \chi_{\pi_{j,j'}}(h) \qquad \text{(II.4.1. 20)}$$

to

$$\delta'(h, 1_{SO(4)}) = \sum_{j} d_{\pi_{j,j}} \chi_{\pi_{j,j}}(h) \qquad \text{(II.4.1. 21)}$$

resulting in

$$Z_P(T) = \int_{\overline{\mathcal{A}}_T} d\mu_{0T}(A) \prod_{\alpha \in \pi_1'(T)} \delta'(A(\alpha), 1_G) \qquad \text{(II.4.1. 22)}$$

(Some version of) (II.4.1. 22) is referred to as the Barrett-Crane model [239]. The model has been improved in its degree of uniqueness by Reisenberger [240] and also by Yetter, Barrett and Barrett and Williams [241].

In contrast to (II.4.1. 18) the integral (II.4.1. 22) is expectedly no longer independent of the triangulation T so that one has to sum over all triangulations in order to obtain triangulation independence. This amounts to defining

$$Z_P = \sum_T w(T) Z_P(T) \qquad \text{(II.4.1. 23)}$$

Of course, the immediate question is how the weight factors $w(T)$ should be chosen. Notice that for this section we mean by a triangulation not an embedded triangulation but a topological one, that is, in some sense four-dimensional diffeomorphism invariance is *defined* to be taken care of.

A clue for how to do that comes from the matrix model approach to two-dimensional quantum gravity, see e.g. [242] and references therein. Boulatov and Ooguri [243] respectively have shown that a Feynman like expansion of a certain field theory over a group manifold (rather than a space time) gives rise to all possible triangulations of the Ponzano Regge (or the Turarev-Viro) model in three dimensions with $G = SU(2)$ and the Crane-Yetter model in four dimensions respectively [231] with $G = SO(4)$. In [244] de Pietri, Freidel, Krasnov and Rovelli applied these ideas in order to recover the Barrett Crane model from a field theory formulation. To see how this works, consider first the case of the BF theory in $D = 3$. Here one considers a real scalar field over $SO(4)^4$ which is right invariant, that is $\phi(h_1, h_2, h_3, h_4) = \phi(h_1 g, h_2 g, h_3 g, h_4 g)$ for any $g \in SO(4)$. One can always obtain such a ϕ from a non-invariant field ϕ' by $\phi = \int_{SU(2)} d\mu_H(g) R_g^* \phi'$. The Boutalov – Ooguri action is then given by

$$\begin{aligned} S'_{BO} &= \int_{SO(4)^4} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) \phi^2(h_1, h_2, h_3, h_4) \qquad \text{(II.4.1. 24)} \\ &\quad + \frac{\lambda}{5!} \int_{SO(4)^{10}} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) d\mu_H(h_5) \times \\ &\quad \times d\mu_H(h_6) d\mu_H(h_7) d\mu_H(h_8) d\mu_H(h_9) d\mu_H(h_{10}) \times \\ &\quad \times \phi(h_1, h_2, h_3, h_4) \phi(h_5, h_6, h_7, h_8) \phi(h_7, h_3, h_8, h_9) \phi(h_9, h_6, h_2, h_{10}) \phi(h_{10}, h_8, h_5, h_1) \end{aligned}$$

which looks almost like a $\lambda\phi^5$ theory. One can now develop the usual Feynman rules for this field theory, giving rise to propagators and vertex functions and construct the perturbation theory as an expansion in powers of λ. The result is (for $\lambda = 1$)

$$\int [d\phi] e^{-S_{BO}(\phi)} = \sum_T w(T) Z_{BF}(T) \tag{II.4.1. 25}$$

with specific weight factors $w(T)$. Notice that the sum over triangulations is redundant for BF theory but not for general relativity.

Given the fact that the Barrett - Crane model basically reduces the $SO(4) \cong SU(2)_L \times SU(2)_R$ of the BF theory to $SU(2)$ it was natural to try to reduce the Crane – Yetter model to the Barrett – Crane model by requiring separate right invariance under $SU(2)$, that is, $\phi(g_1, g_2, g_3, g_4) = \phi(g_1 h_1, g_2 h_2, g_3 h_3, g_4 h_4)$ for any $h_1, .., h_4 \in SU(2)$. Notice that such a field effectively only lives on $SU(2)^4$ precisely as wanted (more precisely, its Peter&Weyl expansion reduces to simple representations). This can be achieved by means of a projection

$$(P\phi)(g_1, .., g_4) = \int_{SU(2)^4} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) \phi(g_1 h_1, g_2 h_2, g_3 h_3, g_4 h_4) \tag{II.4.1. 26}$$

where we have chosen some internal direction in four dimensional Euclidean space in order to write $SO(4)$ in terms of two copies of $SU(2)$ (to choose a $SU(2)$ subgroup of $SO(4)$). The field $P\phi$ is independent of that direction since it is invariant under simultaneous right action by $SO(4)$ as well. The theory considered in [244] is given by (II.4.1. 24) just that ϕ is replaced by $P\phi$, that is,

$$\begin{aligned} S'_{BC} &= \int_{SO(4)^4} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) (P\phi)^2(h_1, h_2, h_3, h_4) \\ &\quad + \frac{\lambda}{5!} \int_{SO(4)^{10}} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) d\mu_H(h_5) \times \\ &\quad \times d\mu_H(h_6) d\mu_H(h_7) d\mu_H(h_8) d\mu_H(h_9) d\mu_H(h_{10}) \times \\ &\times \; (P\phi)(h_1, h_2, h_3, h_4)(P\phi)(h_5, h_6, h_7, h_8)(P\phi)(h_7, h_3, h_8, h_9) \times \\ &\quad \times (P\phi)(h_9, h_6, h_2, h_{10})(P\phi)(h_{10}, h_8, h_5, h_1) \end{aligned} \tag{II.4.1. 27}$$

It was shown that the resulting Feynman expansion indeed gives rise to a sum over triangulations of the Barrett Crane model.

The individual terms of the resulting series, however, are still divergent. In [245] Rovelli and Perez suggested a slight modification of (II.4.1. 27) by removing the projection in the quadratic term, that is,

$$\begin{aligned} S'_{RP} &= \int_{SO(4)^4} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) \phi^2(h_1, h_2, h_3, h_4) \\ &\quad + \frac{\lambda}{5!} \int_{SO(4)^{10}} d\mu_H(h_1) d\mu_H(h_2) d\mu_H(h_3) d\mu_H(h_4) d\mu_H(h_5) \times \\ &\quad \times d\mu_H(h_6) d\mu_H(h_7) d\mu_H(h_8) d\mu_H(h_9) d\mu_H(h_{10}) \\ &\times \; (P\phi)(h_1, h_2, h_3, h_4)(P\phi)(h_5, h_6, h_7, h_8)(P\phi)(h_7, h_3, h_8, h_9) \times \\ &\quad \times (P\phi)(h_9, h_6, h_2, h_{10})(P\phi)(h_{10}, h_8, h_5, h_1) \end{aligned} \tag{II.4.1. 28}$$

which is free of certain bubble divergences in its Feynman expansion. In [246] Perez proved that the resulting model, which is only a slight variation of the Barrett – Crane model and which effectively only depends on simple representations, is actually *finite* order by order in perturbation theory (triangulation refinement). Of course, this does not show that the series converges but it is anyway a remarkable result that no renormalization is necessary. Besides, in [247] it was demonstrated that any Euclidean spin foam model can be written as a field theory over a compact group manifold.

So far we have only discussed the Euclidean theory. Can we also deal with the Lorentzian case ? In [248] Barrett and Crane modified their Euclidean model to the Lorentzian case. One obstacle is that one now has to deal with the non-compact gauge group $SO(1,3)$ for which all non-trivial unitary representations are infinite dimensional. The unitary representations of the universal covering group $SL(2,\mathbb{C})$ are labelled by a pair $(n,\rho) \in \mathbb{R}_0^+ \times \mathbb{N}_0^+$, quite similar to the case of the universal covering group $SU(2) \times SU(2)$ of $SO(4)$ which are labelled by a pair $(j,j') \in \mathbb{N}_0/2 \times \mathbb{N}_0/2$. For an exhaustive treatment see [249]. Following an analogous procedure that has lead to the constraint $j = j'$ in the Euclidean case we now find that the simplicity constraint leads to $n\rho = 0$, that is, either $n = 0$ or $\rho = 0$. These representations pick an $SL(2,\mathbb{R})$ or $SU(2)$ subgroup within $SL(2,\mathbb{C})$ for $n = 0$ or $\rho = 0$ respectively. To see where this comes from, one notices that the B field of the BF theory essentially becomes, upon canonical quantization, an angular momentum operator and the Casimir operators are given by $C_1 = L_{ij}L^{ij}$, $C_2 = L_{ij}(*L)^{ij}$, the simplicity constraint becomes $C_2 = 0$. In the Euclidean case the spectra are $C_1 = j(j+1) + j'(j'+1)$, $C_2 = j(j+1) - j'(j'+1)$ while in the Lorentzian case they become spectra are $C_1 = [n^2 - \rho^2 - 4]/4$, $C_2 = n\rho/4$. We see that in the Euclidean case the simple representations are "spacelike" representations $C_1 \geq 0$ while the simple representations with $n = 0, \rho = 0$ for the Lorentzian theory are timelike and spacelike respectively. The definition of the $\delta-$distribution becomes now more complicated because there is no Peter&Weyl basis any longer. Rather one has direct integrals and sums respectively for the simple continuous and discrete series of representations respectively and in order to evaluate the state sum amplitudes one must now perform also complicated integrals rather than only discrete sums. In [250] Baez and Barrett proved that nevertheless a large class of these amplitudes are "integrable".

In [251] Perez and Rovelli managed to show that also (a variant of) the Lorentzian Barrett – Crane model can be defined as a field theory on a group manifold including the sum over triangulations again. Basically, what one does is to replace in (II.4.1. 28) the group $SO(4)$ by $SL(2,\mathbb{C})$ while the projection P can now be performed with respect to any of the two subgroups $SL(2,\mathbb{R})$ and $SU(2)$ respectively while the field ϕ is now simultaneously $SL(2,\mathbb{C})$ right invariant. In [251] the choice $SU(2)$ was made in order to define P which is therefore given by (II.4.1. 26) with $g_I \in SO(4)$ replaced by $g_I \in SL(2,\mathbb{C})$, $I = 1,2,3,4$. Finally, in [252] Crane, Perez and Rovelli succeeded in proving, using the results of [250], that the field theory [251] is finite order by order in parturbation theory at least on what they call "regular" triangulations.

This concludes our brief report on the impressive progress that has been made over the last few years in the "spin foam model industry". Let us conclude with a couple of remarks:

i) *Spin Foams and Canonical Theory*
What is missing is an interpretation of these spin foam models. Roughly speaking, what one should do is to impose boundary conditions (boundary spin nets) on the partition function and to sum over all spin foam amplitudes and triangulations in between that are compatible with the boundary spin-net. Provided that one can show that the resulting object defines a positive semi-definite sesqui-linear form one can compute its null space and complete the corresponding factor space in order to obtain an inner product. What one then would still have to show is that the theory that one gets implements (some version of) the Hamiltonian constraint. In other words, to be really convincing one must make contact with the canonical theory somehow. An analysis of this kind has been started in [253].

One can try to go the other way around and start from the canonical theory and derive the path integral formulation through some kind of Feynman-Kac formula. A natural starting point for such an analysis would be by using coherent states as has often been stressed by Klauder [216].

ii) *Semiclassical Analysis*
The Perez – Rovelli variant of the Barrett – Crane model seems to be preferred at the moment but it is unclear whether the modification they performed changes the physics significantly or not. Moreover, in some sense there is always a jump in passing from the BF theory to general relativity, in other words, while it is extremely convincing that one should pass to simple representations it would be nicer to start from the constrained BF theory partition function (II.4.1. 15) and arrive at the Barrett – Crane model by integrating over the Freidel – Krasnov – Puzio Lagrange multiplicator. Of course even then one has to make some guesses like the choice of the measure $[dA\ dB\ d\Phi]$. So what one would like to have are some independent arguments that the models proposed have the correct classical limit for instance by showing that they are a well-defined version of the Reisenberger – Rovelli projector (II.4.1. 5).

iii) *Sum over Triangulations*
While we seem to have finiteness proofs for the field theory formulation order by order ("triangulation by triangulation"), it would certainly be even better if one could establish that the sum over triangulations converges. "But maybe this does not need to be the case at all[4]". The reason is that what we really would like to show is that

$$< O >:= \frac{\int[d\phi]e^{-S[\phi]}O(\phi)}{Z} \tag{II.4.1. 29}$$

converges for a sufficiently large set of observables (how to express observables of general relativity in terms of the field theory on the group manifold is another question). This object should be regulated by cutting off the sum over triangulations and then one takes the regulator away. The objects (II.4.1. 29) possibly define the finite moments of a rigorously defined measure on some field space on which the the field ϕ lives. This is exactly how one usually performs constructive quantum field theory, see [32, 122, 125, 128]: Even in free scalar quantum field theory none of the objects $[d\phi], e^{-S[\phi]}, Z$ makes sense separately, it is only the combination $\frac{[d\phi]e^{-S[\phi]}}{Z}$ which can be given a rigorous meaning.

iv) *Built in Causality and Appearance of Renormalization Group*
In dealing with Lorentzian spin foams it is a valid question in which sense the corresponding quantum evolution is causal in any sense. These questions were first addressed in [254] by Markopoulou and Smolin. One may even restrict the class of spin foams to be considered by allowing only those which are causal.

A different question related to the isssue of the classical limit is whether there is some notion of a renormalization group within spin foam models which then would answer the question in which sense they depend on the class of triangulations that we sum over or whether we are allowed to perform small changes in the "initial field theory action" without changing the effective low energy (semiclassical) theory, in other words whether there is a natural notion of universility classes and the like. A first pioneering work has recently been published by Markopoulou [255] in which the Hopf algebra structure underlying renormalization in ordinary

[4]Remark by the author to Alexandro Perez at the "Bleibtreu Meeting", 6th floor, Bleibtreustrasse 12A, 10623 Berlin, Germany, Feb. 16 – 18, 2001, Fotini Markopoulou and Lee Smolin (Organizers).

field theory discovered by Connes and Kreimer [256] was applied to coarsening processes of the triangulations that underly of spin foams.

II.4.2 Quantum Black Hole Physics

A first challenge of quantum black hole physics is to give a microscopic explanation for the Bekenstein – Hawking entropy of a black hole [257] given by

$$S_{BH} = \frac{\mathrm{Ar}(H)}{4\ell_p^2} \tag{II.4.2. 1}$$

where $\mathrm{Ar}(H)$ denotes the area of the event horizon H as measured by the metric that describes the corresponding black hole space time and in this section we set $\ell_p^2 = \hbar G_{Newton}$ instead of $\hbar\kappa = 8\pi\hbar G_{Newton}$.

In [258] Krasnov performed a bold computation: Given any surface S with spherical topology, given some area A and an interval $[A-\Delta A, A+\Delta A]$, let us compute the number N of spin-network states T_s such that $< T_s, \widehat{\mathrm{Ar}}(S)T_s > \in [A-\Delta A, A+\Delta A]$. Of course, N is infinite. But now let us mod out by the gauge motions generated by the constraints: Most of the divergence of N stems from the fact that for given number of punctures $S\cap\gamma(s)$ and fixed representations $\vec{\pi}(s)$, there are uncountably many different spin network states with the same area expectation value because different positions of the punctures give different spin-network states. This is no longer the case after moding out by spatial diffemorphisms. There is, however, still a source of divergence because what matters for the area eigenvalue is more or less only the number of punctures and the spins of the edges that intersect the surfaces S, what happens outside or inside the surface is irrelevant and certainly even after moding by spatial diffeomorphisms one still had $N=\infty$ therefore. Krasnov had to assume that this divergence would be taken care of after moding out the action of the Hamiltonian constraint. Hence, ignoring this final divergence his result for $\Delta \approx \ell_p^2$ was very close to (II.4.2. 1) namely proportional to $\mathrm{Ar}(S)/(4\ell_p^2)$. A similar computation by Rovelli [259] confirmed this value.

This result was promising enough in order to spend more effort in making it water-tight: For instance, nothing in [258] could prevent one from performing the computation for *any* surface, not necessarily a black hole event horizon so that it was conceptually unclear what the computation showed. Somehow one had to invoke the information that H *is* an event horizon into the computation to get rid of the divergences that were just mentioned. Also, given the local nature of the area eigenvalue counting, it was desirable to localize the notion of an event horizon which can be determined only when one knows the entire spacetime (recall that an event horizon [47] is the external boundary of the portion of spacetime that does not lie in the past of null future infinity) which is completely unphysical from an operational point of view because one would never know if a horizon is really an event horizon since the object under study could collide with a burnt out star in the late period of the universe when all life has deceased. Whether or not H is a horizon one should be able to determine by performing local measurements in spacetime.

These questions gave rise to a whole industry of its own, called "isolated horizons", which to a large extent is a new beautiful chapter in classical general relativity. In what follows we will try to summarize the main ingredients of the framework, focussing on the quantum aspects.

Notice that in canonical quantum gravity, as presently formulated, we must specify a three manifold σ of arbitary but fixed topology. When σ has a boundary, one must impose suitable boundary conditions on the fields in order to obtain a well-defined action principle. The idea is to *first* classically encode the presence of a locally defined horizon in the topology of σ and the boundary conditions on

(A, E) at the internal boundary (and the usual asymptotically flat boundary conditions at spatial infinity i^0) and *then* to quantize the system. Let us first give the abstract definition.

Definition II.4.1
A submanifold Δ of a spacetime (M, g) is said to be an isolated horizon if
i)
Δ is topologically $\mathbb{R} \times S^2$, null with zero shear and expansion. This conditions ensures that the covariant derivative ∇ on M induces a unique covariant derivative on Δ via $Du = [\nabla\tilde{u}]_{|\Delta}$ where $\tilde{u}$ is any smooth extension of the vector field u on Δ to M.
ii)
There exists a null normal l of Δ such that $(\mathcal{L}_l D - D\mathcal{L}_l)u \sim 0$ where $\sim$ denotes equality when restricted to Δ (since l is defined on M we must use an extension $\tilde{u}\tilde{Du}$ of u, Du in order to act with the Lie derivative $\mathcal{L}_l$).
iii)
The field equations hold at Δ.

Notice that this definition is local to Δ. We can think of l as "time-direction" on Δ and so the first two conditions imply that the geometry on Δ is stationary with respect to l. These three conditions are very tight but less tight than those that lead to event horizons although all known black hole families are encompassed. For instance it allows that there is radiation within the bulk of M which may even fall into the singularity as long as it does not cross Δ. For a brief discussion of all implications and an extension to matter see [228] and for a detailed derivation see [230].

For our limited considerations concerning the black hole entropy calculation it will be sufficient to describe the consequences of definition II.4.1 for the canonical quantization. From now on we will restrict our attention to the portion of M which is bounded by two initial data hypersurfaces Σ_1, Σ_2, spatial infinity i^0 and the isolated horizon Δ. As it is clear from the definition, the isolated horizon implies that our initial data hypersurfaces Σ that foliate M are diffeomorphic to σ where σ has an internal S^2 boundary. Then definition II.4.1 implies the following (we will work with arbitrary Immirzi parameter β but suppress it in the canonical coordinates $(A = \Gamma + \beta K, E = E_1/\beta)$):
1)
There is a differentiable bijection $r^j : S^2 \to S^2$ (meaning that $r^j r^k \delta_{jk} = 1$). Given an $SU(2)$ principal fibre bundle over σ we obtain a principal $U(1)$ bundle over $S := S^2$ by restricing the fibres over $s \in S^2 \subset \sigma$ to those $g \in SU(2)$ which preserve the internal vector r^j. Consider the $U(1)$ connection on S^2 defined by

$$W := -\frac{1}{\sqrt{2}}[X^*\Gamma^j]r_j \qquad \text{(II.4.2. 2)}$$

where $X : S \to \sigma$ denotes the corresponding embedding and Γ the spin connection of the triad.
2)
The boundary conditions on the pull-backs $\underline{A}, \underline{*E}$ to S^2 are that

$$\begin{aligned} \underline{A}^j &:= X^*A^j = Wr^j \\ \underline{*E} &:= X^*(*E^j)r_j = -\frac{a_0}{2\pi\beta}dW \end{aligned} \qquad \text{(II.4.2. 3)}$$

where $a_0 := \mathrm{Ar}_E(S)$ is constrained to be a constant by the isolated horizon conditions (independent of E). Thus $(\underline{A}, \underline{*E})$ are completely determined by (W, dW) respectively.
3)
Finally, the symplectic structure of our classical system turns out to be

$$\Omega((\delta A, \delta E), (\delta A, \delta E)) = \frac{1}{\kappa}[\int_{\sigma - S} \mathrm{Tr}(\delta A \wedge \delta * E' - \delta A' \wedge \delta * E) + \frac{a_0}{\pi\beta}\int_S \delta W \wedge \delta W'] \qquad \text{(II.4.2. 4)}$$

for arbitrary tangential vectors $\delta A, \delta E, \delta A', \delta E'$ to the phase space $\mathcal{M}$.

Thus $\Omega = \Omega_\sigma + \Omega_S$ consists of a bulk and a surface term. Clearly, classically the surface degrees of freedom are determined by the bulk degrees of freedom by continuity but this will change in quantum theory where the distributional nature of the quantum configuration space excites additional degrees of freedom. Interestingly, the surface symplectic structure is that of a $U(1)$ Chern-Simons theory on $\mathbb{R} \times S$ with action

$$S_{CS} = \int_{\mathbb{R}\times S} W \wedge dW = \int_{\mathbb{R}} dt \int_S \epsilon^{IJ}[\dot{W}_I W_J + W_t(dW)_{IJ}] \qquad \text{(II.4.2. 5)}$$

which displays $\epsilon^{IJ} W_J$ as the momentum canonically conjugate to W_I and the constraint is that W be flat.

In order to quantize the system we will adopt the following strategy:

1.) *Quantum Configuration Space*
Essentially we will make the bulk and surface configuration degrees of freedom independent of each other, that is, $\overline{\mathcal{A}} = \overline{\mathcal{A}}_\sigma \times \overline{\mathcal{A}}_S$ with distributional $SU(2)$ and $U(1)$ connections respectively.

2.) *Kinematical Hilbert Space*
Accordingly the kinematical Hilbert space adopts a tensor product structure $\mathcal{H}^0 = \mathcal{H}^0_\sigma \otimes \mathcal{H}^0_S$ with $\mathcal{H}^0_\sigma = L_2(\overline{\mathcal{A}}_\sigma, d\mu_0), \mathcal{H}^0_S = L_2(\overline{\mathcal{A}}_S, d\mu_0)$.

3.) *Quantum Boundary Conditions*
This structure suggests to solve (II.4.2. 3) in the symbolic form

$$[1_{\mathcal{H}^0_\sigma} \otimes e^{i\widehat{dW}}]\psi = [e^{-\frac{2\pi\beta i}{a_0}\widehat{\underline{*E}}} \otimes 1_{\mathcal{H}^0_S}]\psi \qquad \text{(II.4.2. 6)}$$

The reason for this particular exponentiation of (II.4.2. 3) is required by the particulars of the quantization of Chern-Simons theory.

4.) *Implementation of Quantum Dynamics*
Finally one has to impose the constraints at S. It turns out that $\widehat{\underline{*E}}$ and $\widehat{dW}$ generate $U(1)$ gauge transformations in the bulk close to S and on S respectively so that through (II.4.2. 6) *the Gauss constraint is already satisfied*, in other words, the total state depending on both bulk and surface degrees of freedom is gauge invariant! Next, as already anticipated at the beginning of this section, the diffeomorphism constraint restricted to S basically tells us that what is important is the number of punctures of the graph of a spin-network with the surface S and not their position. Finally, the Hamiltonian constraint vanishes identically at S due to the definition of an isolated horizon which, in particular, requires that lapse functions that generate gauge motions induced by H must be identically zero at S (This does not mean that the lapse of a classical isolated horizon solution must vanish at S, rather there is a subtle difference between lapse functions that generate symmetries rather than gauge transformations (see, e.g. [87] and references therein) so that in this case lapse functions that do not vanish at S map between gauge inequivalent solutions).

Let us describe this in more detail. In order to solve (II.4.2. 6) one makes a tensor product ansatz $\psi = \psi_\sigma \otimes \psi_S$ implying

$$\psi_\sigma \otimes [e^{i\widehat{dW}}\psi_S] = [e^{-\frac{2\pi\beta i}{a_0}\widehat{\underline{*E}}}\psi_\sigma] \otimes \psi_S \qquad \text{(II.4.2. 7)}$$

and one will try to look for eigenvectors of the exponentiated operators with the same eigenvalues. Now, what we have quantized in section I.3 was not $*E^j$ bat rather $\int_{S'} *E$ where S' is any surface, however, the fact that this function can be turned into an operator means that there exists an operator valued distribution which when restricted to S is given on a function cylindrical over γ by

$$\widehat{\underline{*E}}(s)\psi_{\sigma,\gamma} = 4\pi i\ell_p^2 \sum_{p\in S\cap\gamma} \eta\delta^{(2)}(s, X^{-1}(p)) \sum_{e\in E(\gamma);f(e)=p} L_e^j r_j(s)\psi_{\sigma,\gamma} \tag{II.4.2. 8}$$

(the appearance of the 4π rather than $1/2$ as compared to section I.3 is due to our definition of ℓ_p in this section). Here $\eta = \frac{1}{2}\epsilon_{IJ} ds^I \wedge ds^J$ in a local system of coordinates s^I on S. Notice that only left invariant vector fields appear because every edge that intersects S is of the "down" type (S carries outward orientation and there is no interior of σ with respect to S because S is a boundary). No edge of the bulk graph γ lies inside S because edges inside S label surface degrees of freedom. Now the spectrum of (II.4.2. 8) can be computed by inspection: The operator ir_jL^j on $L_2(G, d\mu_H)$ is nothing else than the operator $2\hat{J}_3$ for the quantum mechanics of the angular momentum whence it has spectrum $2m$ with m a half-integral quantum number. In fact, we can choose a spin network basis in which $(L_e^j)^2, L_e^j r_j(s)$ are diagonal and immediately obtain as "distributional eigenvalues"

$$\widehat{\underline{*E}}(s)\psi_{\sigma,s} = 8\pi\ell_p^2 \sum_{p\in S\cap\gamma(s)} \eta\delta^{(2)}(s, X^{-1}(p)) \sum_{e\in E(\gamma(s));f(e)=p} m_e\psi_{\sigma,s} \tag{II.4.2. 9}$$

where $|m_e| \leq j_e$ is half-integral. The result (II.4.2. 8) motivates to split the bulk Hilbert space as

$$\mathcal{H}_\sigma^0 = \oplus_{\mathcal{P},m}\mathcal{H}_\sigma^{\mathcal{P},m} \tag{II.4.2. 10}$$

where $\mathcal{P}$ denotes the set of all punctures of S (that is, positions of points where a bulk graph intersects S) and m the possible eigenvalues (II.4.2. 8).

Next we turn to the operator $\widehat{dW}$. Since its eigenvalues must match (II.4.2. 8) we conclude that the quantum curvature of W is flat everywhere except at the punctures. Consider the spaces $\mathcal{A}^\mathcal{P}, \mathcal{G}^\mathcal{P}, \mathcal{D}^\mathcal{P}$ of connections which are flat up to the punctures $\mathcal{P}$, gauge transformations which equal the identity at $\mathcal{P}$ and analytic diffeomorphisms which preserve $\mathcal{P}$. Consider the moduli space

$$\mathcal{M}^\mathcal{P} := \mathcal{A}^\mathcal{P}/(\mathcal{G}^\mathcal{P} \triangleleft \mathcal{D}^\mathcal{P}) \tag{II.4.2. 11}$$

and turn it into a symplectic manifold by equipping it with the Chern-Simons symplectic structure (the semi-direct product in (II.4.2. 11) appears because diffeomorphisms act non-trivially on gauge transformations). The phase space $\mathcal{M}^\mathcal{P}$ is compact (one way to see this is that it can be coordinatized by $U(1)$ holonomies, see below) and therefore does not admit the standard cotangent bundle polarization. However, it can be quantized by the methods of geometric quantization [51] by choosing a (positive) Kähler polarization. It would take us too far to develop the necessary background for general geometric quantization and quantization of Chern-Simons theory in particular, see however the exhaustive treatment in the beautiful thesis [257]. The outcome of this analysis is as follows:

1) *Phase Space*

First of all, $\mathcal{M}^\mathcal{P}$ can be identified with the torus with $T^{2(n-1)} := \mathbb{C}^{n-1}/(2\pi\mathbb{Z})^{2(n-1)}$ where $n = |\mathcal{P}|$ is the number of punctures. To see at least intuitively how this happens, notice that the holonomies around loops and paths between the punctures separate the points of $\mathcal{A}^\mathcal{P}/\mathcal{G}^\mathcal{P}$ since gauge transformations at the punctures are trivial. The punctured surface S is homeomorphic to a sphere with holes and the holonomy of a flat connection along a loop depends only on its homotopy type. Thus, one might think that the homotopy group of of the punctured sphere is generated by n elements α_p where α_p encloses p but not any other puncture, however, this is not true: Any loop β which encloses all punctures is contractible "over the back of the sphere" and so is the loop $\beta \circ (\circ_{p\in\mathcal{P}}\alpha_p^{-1})$. Thus we may get rid of one of the α_p, say α_{p_0} for some fixed puncture p_0. Next, consider paths $e_p,\ p \neq p_0$ between p and p_0 which do not intersect any $\alpha_{p'},\ p' \neq p, p_0$. Obviously, any path between $p, p' \in \mathcal{P}$ is homotop to $e_p \circ e_{p'}^{-1}$. This explains already why the phase space should have dimension $2(n-1)$. But each holonomy takes values in $U(1)$ which is diffeomorphic with $S^1 = T^1 = \mathbb{R}/(2\pi\mathbb{Z})$. Finally, the diffeomorphisms in $\mathcal{D}^{\mathcal{P}}$ preserve the homotopy type of the $\alpha_p, e_p;\ p \neq p_0$.

2) *Geometric Quantization*
Let us introduce coordinates $x_p, y_p \in \mathbb{R}/(2\pi\mathbb{Z})$ with $h_{\alpha_p} = e^{ix_p}, h_{e_p} = e^{iy_p}$. The numbers x_p, y_p coordinatize a point in $\mathcal{M}^{\mathcal{P}}$ and we can always find a representative

$$W = \sum_{p \neq p_0} [x_p X_p + y_p Y_p] \qquad \text{(II.4.2. 12)}$$

where the one forms X_p, Y_p are the Poincaré duals to the e_p, α_p, that is

$$\int_{e_p} X_{p'} = \int_{\alpha_p} Y_{p'} = \delta_{pp'} \text{ and } \int_{\alpha_p} X_{p'} = \int_{e_p} Y_{p'} = 0 \qquad \text{(II.4.2. 13)}$$

Pulling back the symplectic structure Ω_s by (II.4.2. 13 we find

$$\Omega = \frac{k}{2\pi} \sum_{p \neq p_0} dx_p \wedge y_p \qquad \text{(II.4.2. 14)}$$

Now the detailed framework of geometric quantization reveals that there is an obstruction to quantization which is spelled out in terms of Weil's integrality criterion [51]. In our case it boils down to the condition that the so-called level of the Chern-Simons theory

$$k := \frac{a_0}{4\pi\beta\ell_p^2} \qquad \text{(II.4.2. 15)}$$

must be an integer.

Next consider holomorphic wave functions of the $z_p = x_p + iy_p$ which defines a so-called Kähler polarization (similar to the Segal-Bargmann representation for the phase space $\mathbb{R}^{2(n-1)}$). We can view functions on $T^{2(n-1)}$ as functions on $\mathbb{C}^{2(n-1)}$ which are invariant under translations within the lattice $\Lambda = (2\pi\mathbb{Z})^{2(n-1)}$, that is, periodic *holomorphic* functions. Now by Liouville's theorem periodic holomorphic functions do not exist so the best that one can achieve are quasi-periodic functions which are also called $\Theta-$functions [261]. These are holomorphic functions which depend on the Teichmüller parameter τ with $\Im(\tau) > 0$ which determines the complex structure of the torus and the positive level integer $k > 0$. They satisfy the functional equations (in one complex dimension) $\Theta^k_\tau(z + 2\pi) = \Theta^k_\tau(z)$, $\Theta^k_\tau(z + 2\pi\tau) = \exp(-ikz + b)\Theta^k_\tau(z)$ where b is an arbitrary complex number. It turns out that the vector space of functions satisfying these functional equations is real $k-$dimensional so that we get k solutions $\Theta^k_\tau(z, a)$ with $a = 0, 1, .., k-1 \in \mathbb{Z}_k$. In our case we have $\tau = i$. The significance of the level k is that the symplectic structure depends on it and that the $\Theta-$functions of level k determine a $k-$dimensional representation of the Heisenberg group generated by the exponentials of the (pre-)quantum operators $\hat{x}_p, \hat{y}_p$. The final result is that in our case only Θ functions $\psi_{S,a}$ labelled by $\vec{a} \in (\mathbb{Z}_k)^{n-1}$ are indistinguishable, in fact, they form a basis in the prequantum Hilbert space of square integrable (with respect to the Liouville measure times a damping factor related to the Kähler potential) holomorphic sections (of a complex line bundle over the phase space). (It may come as a surprise that therefore $\mathcal{H}^{\mathcal{P}}_S$ is finite dimensional, namely k^{n-1} but it really is not because the number of quantum degrees of freedom is roughly given by the Liouville volume of the phase space (which in our case is finite) divided by the volume of a phase cell). For the same reason the holonomy operators $\hat{h}_{\alpha_p}$ have eigenvalues $e^{2\pi i a_p/k}$. The operators $\hat{h}_{e_p}$ will disappear from the final picture since we have to take the quotient later on also with respect to gauge transformations which are not trivial at the punctures.

We thus conclude that the geometric quantization of $\mathcal{M}^{\mathcal{P}}$ leads to a Hilbert space $\mathcal{H}_S$ which is given by the inductive limit of the Hilbert spaces $\mathcal{H}_S^{\mathcal{P}}$ where $\mathcal{P}$ ranges over all finite point subsets of S and where $\mathcal{H}_S^{\mathcal{P}}$ is isomorphic with the geometric quantization of the corresponding torus. On $\mathcal{H}_S^{\mathcal{P}}$ the holonomy operators $\hat{h}_p$, can be simultaneously diagonalized these operators and their eigenvalues are given by $\hat{h}_p \psi_{S,a} = e^{2\pi i a_p/k}$ where $\sum_{p\in\mathcal{P}} a_p = 0(\text{mod } k)$.

Let now $S_p \subset S$ be the interior of α_p then the non-distributional way to state (II.4.2. 7) is given by

$$\psi_\sigma \otimes [\hat{h}_p \psi_S] = [e^{i - \frac{2\pi\beta i}{a_0}\int_{S_p} \widehat{\underline{*E}}} \psi_\sigma] \otimes \psi_S \qquad \text{(II.4.2. 16)}$$

for any $p \in \mathcal{P},\ \psi_S \in \mathcal{H}_S^{\mathcal{P}},\ \psi_\sigma \in \mathcal{H}_\sigma^{\mathcal{P}}$. This evidently leads to the condition

$$m_p = -a_p(\text{mod } k) \qquad \text{(II.4.2. 17)}$$

for the corresponding eigenvalues and we conclude that the kinematical Hilbert space is given (modulo completion) by

$$\mathcal{H}^0 = \oplus_{\mathcal{P},m,a;\, 2m=-a(\text{mod } k)} \mathcal{H}_\sigma^{\mathcal{P},m} \otimes \mathcal{H}_S^{\mathcal{P},a} \qquad \text{(II.4.2. 18)}$$

This Hilbert space is easily seen to solve the full Gauss constraint already since we require states to be gauge invariant in the bulk away from S and the condition (II.4.2. 13) is exactly the gauge invariance condition for gauge transformations at the punctures (actually with respect to a necessarily reduced gauge group).

Next we have to reduce with respect to the spatial Diffeomorphism constraint which simply amounts to replacing the spaces $\mathcal{H}_\sigma^{\mathcal{P}}, \mathcal{H}_S^{\mathcal{P}}$ by $\mathcal{H}_\sigma^n, \mathcal{H}_S^n$ where now only the number of punctures is relevant.

Finally, with respect to the Hamiltonian constraint there is nothing left to do for the reason already mentioned above which is quite lucky because, as we have said before, there is no proof that the proposed bulk Hamiltonian constraint is the correct one but whatever it is it will vanish at S.

Our final task will be to do the entropy counting. The isolated horizon area operator $\widehat{\text{Ar}}(S)$ is a true Dirac observable in the present situation since it is gauge invariant by construction and diffeomorphism invariant under $\text{Diff}(S)$. Its eigenvalues on the sector of the physical Hilbert space labelled by $(n, \vec{j}, \vec{m}, \vec{a})$, $2\vec{m} + \vec{a} = 0$ (mod k) with $|m_l| \le j_l$ the spin of the edge entering the l'th puncture (recall that $(L_l^j)^2$ and $L_l^j r_j$ can be diagonalized simultaneously with eigenvalues $4j_l(j_l+1)$ and $2m_l$ respectvely) are given by

$$\text{Ar}(n, \vec{j}) = 8\pi\ell_p^2\beta \sum_{l=1}^n \sqrt{j_l(j_l+1)} \qquad \text{(II.4.2. 19)}$$

(the area operator acts on the bulk degrees of freedom only). We are looking now for all those sectors for which (II.4.2. 19) lies in the interval $[a_0 - \delta, a_0 + \delta]$. Notice that when $n, \vec{j}, \vec{m}$ are given then $\vec{a}$ is completely fixed already. The crucial point is now that

$$\mathcal{H}_\sigma^{n,\vec{j},\vec{m}} = \otimes = \mathcal{H}_V^{n,\vec{j},\vec{m}} \otimes \mathcal{H}_{bh}^{n,\vec{j},\vec{m}} \qquad \text{(II.4.2. 20)}$$

where the first factor (corresponding to edges not intersecting S) has infinite dimension while the second has finite dimension (corresponding to edges intersecting S). We can summarize this in the formula (modulo completion)

$$\mathcal{H}_{phys} = \oplus_{n,\vec{j},\vec{m},\vec{a}=_k-2\vec{m}} \mathcal{H}_V^{n,\vec{j},\vec{m}} \otimes \mathcal{H}_{bh}^{n,\vec{j},\vec{m}} \otimes \mathcal{H}_S^{n,\vec{a}} \tag{II.4.2. 21}$$

Next we form the corresponding microcanonical statistical ensemble with the density matrix

$$\hat{\rho}_{bh} := 1_V \otimes [\frac{1}{N_{a_0,\delta}} \sum_{n,\vec{j},\vec{m}} |\psi_{bh}^{n,\vec{j},\vec{m}} >< \psi_{bh}^{n,\vec{j},\vec{m}}|] \otimes 1_S \tag{II.4.2. 22}$$

where the sum is over all black hole sectors compatible with a_0, δ and $N_{a_0,\delta}$ is their number. Clearly, the quantum statistical entropy is given by

$$S = -\mathrm{Tr}(\hat{\rho}_{bh} \ln(\hat{\rho}_{bh})) = \ln(N_{a_0,\delta}) \tag{II.4.2. 23}$$

Thus we just need to count states. The analysis is not entirely straightforward but can be summarized as follows:
It turns out that expectedly most of the entropy comes from those configurations with $j_l = 1/2$ (maximum disorder). Then $\mathrm{Ar}(n,\vec{j}) = 4\pi\ell_p^2\beta n\sqrt{3}$. If we choose $\delta > 8\pi\ell_p^2\beta\sqrt{3}$ then we always find an even integer n in order that the eigenvalue lies in the required interval. Now the eigenvalue $j_l = 1/2$ has degeneracy 2 corresponding to the two possible projections $m_l = \pm 1/2$ suggesting that there are 2^n such states, *one Boolean degree of freedom per puncture.* However, we must satisfy the constraint $2(m_1 + .. + m_n) = 0(\mathrm{mod}\ k)$. Certainly for large a_0 we have $|2(m_1 + .. + m_n)| \leq n \approx k/\sqrt{3} < k$ which means that actually $m_1 + .. + m_n = 0$, thus half of the spins must be up the others are down. This brings the number of states with $j_l = 1/2$ down to $\begin{pmatrix} n \\ n/2 \end{pmatrix}$ which therefore is a lower bound for the number $N_{a_0,\delta}$. The derivation of an upper bound is more complicated but can be done with the result that the leading order term is given by $S = \ln(2)n_0$ where $n_0 \approx k/\sqrt{3}$ which would already be the leading order if we just had taken the lower bound and applied Stirling's formula. We see that we *precisely* reproduce the Bekenstein-Hawking entropy provided we choose the Immirzi parameter to be

$$\beta = \frac{\ln(2)}{\pi\sqrt{3}} \tag{II.4.2. 24}$$

Each of these three transcendent numbers has a well-understood origin: $\pi = \kappa/(8G_{Newton})$, $\sqrt{3}/2$ is the lowest non-vanishing eigenvalue of $\sqrt{j(j+1)}$ and $\ln(2)$ comes from $\ln(2^n)$.

The strategy to choose the Immirzi parameter according to (II.4.2. 24) would be worthless if it would not be the *same* value that one would have to match for various kinds of black holes, not only the vacuum black holes that we have treated so far. However, as one can show [262] even for dilatonic and Yang-Mills hair black holes the same value works. This relies on the following facts: a) the presence of this bosonic matter does not change the isolated horizon boundary conditions, b) the matter fields are determined through W at S and therefore c) matter has no independent surface degrees of freedom. It should be pointed out that all of this works for astrophysically realistic (Schwarzschild), four-dimensional, non-supersymmetric black holes.

A couple of remarks are in order:

i) *Non-Triviality*

The calculation is certainly very impressive because one could not have expected from the outset that there would be a harmonic interplay between classical general relativity (isolated horizon boundary conditions), quantum gravity (discrete eigenvalues of the area operator) and quantum Chern-Simons theory (horizon degrees of freedom).

ii) *Other Results*
Recall that there has been established a precise dictionary between the four laws of usual thermodynamics and black hole thermodynamics for *event horizons*. It turns out that one can write another dictionary for *isolated horizons* [263]. Also cosmological horizons can be described by isolated horizon methods.

Next, we have pointed out before that the main series of the area spectrum is by far not all of it. In particular, one can show [265] that the number of eigenvalues in the interval $[a_0 - \ell_p^2, a_0 + \ell_p^2]$ grows as $e^{\sqrt{a_0}/\ell_p}$ explicitly ruling out the naive ansatz made in [266] that the area spectrum is evenly spaced which seemed to be supported by the errornous computation [161] and would have implied an even spacing of the spectrum. This has huge observational consequences: The peak of the black body Hawking spectrum from the black hole is at frequencies $\omega_0 \approx 1/r_0$ where $r_0 \approx GM$ is the Schwarzschild radius of the black hole (we neglect numerical constants and set $c = 1$). Now $a_0 = 4\pi r_0^2$ and since energy emission of the black hole is due to "area transitions" we obtain spectral lines at $\hbar\omega \approx (\Delta M) \approx \Delta(\sqrt{a}/G) \approx (\Delta a)/(gr_0) \approx \omega_0 \Delta a/G$. We see that if the spectrum would be evenly spaced at $\Delta a \approx \hbar G$ then $\omega \approx n\omega_0$ so we would not get a black body spectrum at all, every line would be at a multiple of the peak line.

iii) *Open Problems*
The case that we have treated above was for a static isolated horizon. While rotating isolated horizons can be treated classically [264] so far the quantum theory has not been developed. A related question is whether one can also treat Hawking radiation with the present framework and a pioneering ansatz was made in [267]. Also, it has been conjectured that the Bekenstein Hawking entropy is an inevitable, universal property of any kind of quantum gravity theory and a proof of that conjecture was begun in [268]. However, this calculation was shown not to apply in the present context [269]. Finally, a better understanding about the role of the Immirzi parameter and whether or not it should be fixed as displayed here would be desirable.

iv) *First Principle Calculation*
The isolated horizon description is an effective one (not from first principles) because the presence of an isolated horizon was put in at the classical level. It would be far more desirable to begin with the full quantum theory and to have *quantum criteria* at one's disposal for when a given state represents a quantum black hole. At this point the semi-classical discussed in section II.3.2 could be of some help.

II.4.3 Connections Between Canonical Quantum General Relativity and String – (M) Theory

Smolin has conducted an ambitious programme, namely the investigation of possible interfaces between canonical quantum general relativity and M – Theory, especially in its Matrix theory incarnation [270]. The ultimate goal of this effort is to arrive at a background independent formulation of M – theory.

The possible links between these two major approaches to quantum gravity are very complex and even a brief introduction would require at least some background material for M – Theory which would really go much beyond what this review is intended to cover. We therefore must unfortunately leave the reader with the literature cited and just point out, as an example, the recent paper [271] which seems to indicate that there is a conflict between Maldacena's conjecture [272], which says that superstring theory on an Anti – deSitter (AdS) background spacetime is equivalent ("dual") to a conformal field theory (a super Yang Mills theory) on the conformal boundary of the AdS space in the sense that all scattering amplitudes in the bulk are completely determined by the scattering amplitudes on the boundary, and Rehren's duality [273], which says that there is a natural isomorphism between nets of local algebras (in the sense of the Haag – Kastler formulation of quantum field theory) in the bulk and on the conformal boundary whenever the AdS background is available. This has actually already been observed earlier in [274]. The conflict is the following: Rehren's duality maps a local theory (spacelike separated algebras commute), say the one on the boundary, to a theory which is local, say in the bulk, however, not in the standard way: If the former theory is based on a Lagrangean principle so that fields can be indexed by points on the boundary, the latter theory *does not admit a Lagrangean formulation*, it is in that sense a non-local theory, in particular, there is no causal propagation in the standard sense. Looking at the details of Rehren's map the technical reason for this effect is that one has dropped one dimensional information in this "algebraic holography".

On the other hand, the Maldacena conjecture seems to assume a map between two Lagrangean theories (the effective low energy theory of string theory in the bulk is a supergravity theory). Thus, either the two dualities have nothing to do with each other because (the low energy limit of) string theory on a given background is not a theory to which the Haag-Kastler framework applies (which would be extremely surprising) or there is no Lagrangean origin for M – Theory, not even for its low energy limt (which is contrary to all what string theorists seem to assume for decades).

II.5 Selection of Open Research Problems

Instead of summarizing the summary that we have given in this review, we close with a (far from exhaustive) list of open research problems which the author considers most pressing to be solved.

1) *Spectrum of the Volume Operator*
We have seen that the volume operator plays a prominent role in the dynamical structure of the theory, it enters the Hamiltonian constraint, matter Hamiltonians, the generator of the Wick transform, the Quantum Dirac Algebra, the asymptotic Poincaré algebra and the (possible) complexifier of coherent states. Despite this pivotal interaction of quantum dynamics with the volume operator, very little is known about its spectrum not even asymptotically (that is, in the limit of large spins) and it would therefore be highly desirable to gain more control over it.

2) *Rigorous Construction of the Wick Transform*
For various reasons it would be of benefit to return to a complex connection formulation, if only because it is closer to a manifestly covariant formulation of the theory in terms of a path integral. It therefore seems to be mandatory to construct the generator of that transform as a self-adjoint operator. For the beginning of a corresponding analysis in the simplified context of mini-superspace models see [275]. Related, although independent to this, is the question whether there exist background independent measures on distributional spaces of connections for non-compact gauge groups in analogy to the structure provided by $\overline{\mathcal{A}}, \mu_0$.

3) *Correct Version of the Hamiltonian Constraint Operator*
We have tried to indicate that at this point we do not have *the* Hamiltonian constraint but a huge class of consistent proposals. None of them seems completely satisfactory though for the reasons we have mentioned. It would be worthwhile to explore which kind of generalizations are allowed which still lead to anomaly-free constraint algebras while having the correct classical limit.

4) *Proof of the Correct Classical Limit*
At this point the semi-classical analysis has just started. The semi-classical states that are available all suffer from one and the same desease: They are incapable to reproduce correct expectation values for operators which map between spin network states over different graphs. But this is precisely what both the Hamiltonian and Diffeomorphism constraint do. An appropriate improvement is therefore mandatory before we can even seriously ask the question whether we have the correct theory.

5) *Contact with Quantum Field Theory on Curved Spacetimes*
We have indicated at various occasions how to make contact with quantum field theory on curved spacetimes. Luckily, for bosonic quantum field theories these questions can already be addressed with the semi-classical states available. In its present form, quantum field theory on curved spacetimes are most naturally formulated in the language of algebraic quantum field theory [276]. It is almost clear that nets of local algebras can have at most a semi-classical meaning since the axiom of locality makes sense only when one has a background spacetime available, in other words the fluctuations of the quantum metric must be small compared to those of matter. It would be crucial to make this correspondence manifest.

6) *Three – and Four Dimensional Dirac Observables*
We have indicated in this article how one could use matter in order to turn the area operator

into a spatially diffeomorphism invariant operator. We need something similar with respect to the Hamiltonian constraint as well, at least we should have a constructive procedure for how to arrive at them by some algorithm which converges sufficiently fast so that one has the notion of an approximate Dirac observable at least. See [277] for a first proposal.

7) *Avoidance of Classical Singularities*
As we mentioned, there are indications that the quantum symmetry reduction of quantum general relativity to certain Bianchi cosmologies predicts that there is no big bang singularity at all. It would add faith to these results if one could establish similar results within the full theory without symmetry reduction.

8) *Hawking Effect from First Principles*
As mentioned in the previous section, the isolated horizon approach to quantum black holes, as every approach to quantum black holes that is presently available, needs a classical input. It would be far more convincing if one could develop pure quantum criteria for when a state describes a black hole, thereby opening the possibility to derive the Hawking effect from first principles.

9) *Proof of a Feynman – Kac Formula for Canonical Quantum Gravity*
What we need in order to connect the spin foam approach with the canonical approach is some kind of Feynman – Kac formula, some possibility to derive the spin foam model out of canonical quantum gravity or vice versa. This would be highly desirable since path integrals and operator methods usually complement each other. Again, coherent states could play a crucial role towards this goal.

10) *Combinatorial Formulation of the Theory*
An ugly feature of the present framework is that it still depends in a technically not too weak way on a background differential (or even analytic) manifold and topology. If, as many suspect, quantum gravity should allow for topology change then we must get rid of these structures. Actually, the present framework suggests its own way out of these limitations: Instead of talking about embedded graphs we must learn how to formulate the theory over algebraic graphs, see [138, 278].

11) *Introduce Higher Form Variables in Higher Dimensions*
It is quite possible that supersymmetry does not play any role in nature and that actually four spacetime dimensions are sufficient. However, if M – Theory is correct then quantum gravity in four dimensions can be at most an effective theory. In order to address this possibility one can start by trying to develop a background independent quantization of 11D supergravity, the low energy limit of M – Theory. As we have seen, in order to achieve background independence one must build the canonical theory on $p-$form fields rather than metrics, similar as in four dimensions. For a first ansatz see [279].

12) *Make contact with String (Membrane) Theory* Related to this is the question whether methods of background independent quantum gravity developed in three and four dimensions cannot be used also in higher dimensions, for instance in quantizing the super-membrabe non-perturbatively. The super-membrane in 11D is one of the "hot" candidates for M – Theory.

13) *Construction of Physical Inner Product*
We must develop an algorithm for how to arrive at a physical inner product (automatically incorporating the correct adjointness relations) for open constraint algebras, at least in principle.

14) *Deparameterization, Reconstruction Problem*
Suppose somebody is going to find the complete space of solutions to all quantum constraints, a consistent inner product with all desired properties and a complete set of Dirac observables. What is she/he going to do with it ? By definition "nothing ever happens in quantum gravity" meaning that the Dirac observables are "constants of motion" (strictly speaking only if σ is compact). But although our world is four-dimensionally diffeomorphism invariant (and therefore all true observables should be highly non-local) we perform local measurements every day. We must learn how to recover such a local description from the frozen picture that we are confronted with in quantum gravity. This is the reconstruction problem [222]. For a pretty proposal see [61, 62].

15) *Representation Independent Formulation*
The lesson that we learn from algebraic quantum field theory is that the important, primary ingredient are nets of algebras of operators (which in our case, however, would presumably be rather non-local). Only in a second step one studies representations of these algebras, of course one must make sure that there exist physically interesting ones. The advantage of this purely algebraic approach is that one can perform a representation free structural analysis of the theory. It would, for instance, be important to have an analog of the DHR analysis (Doplicher, Haag, Roberts) for the classification [5] of available representations for our theory at one's disposal in order to know which features are tied to a specific representation and which are not. Likewise it would be worthwhile thinking about a suitable background free generalization of the Haag – Kastler axioms. For some first steps in that direction, although within the constructive approach, see [128].

There is an endless chain of other problems in quantum general relativity both on the technical and on the conceptual side which we cannot possibly enumerate here. Hopefully, one day bright students will figure them all out.

Part III

Mathematical Tools

III.1 The Dirac Algorithm for Field Theories with Constraints

It is a crime that the subsequent analysis is not a standard ingredient of every course in theoretical mechanics. Every interaction that we know today underlies a gauge theory, that is, a field theory with constraints. However, constraints are generically at most mentioned and one usually finds out about the fact that one was truly betrayed in that beginning theoretical mechanics course only much later. This is the more astonishing as this really important topic can be taught at a truly elementary level. Also quantum mechanics is not needed (at most for motivational purposes), the theory can be formulated in purely classical terms. We recommend the classic expositions by Dirac [24] and by Hanson et al [25] as introductory texts. More advanced are the textbooks [200] and [26]. For geometrical quantization with constraints see [27] and for a more mathematical formulation see [28].

We will consider only a finite number of degrees of freedom. The more general case can be treated straightforwardly, at least at a formal level. We will also not consider the most general actions but only those which lead to phase spaces with a cotangential bundle topology. For the more general cases see the cited literature.

Definition III.1.1
Consider a Lagrangean function $L:\; T_*(\mathcal{C}) \to \mathbb{C};\; (q^a, v^a) \mapsto L(q,v)$ on the tangential bundle over the configuration manifold $\mathcal{C}$ where $v := \dot{q}$ (velocity) defines the corresponding action principle.
i)
The map

$$\rho_L:\; T_*(\mathcal{C}) \to T^*(\mathcal{C});\; (q,v) \mapsto (q, p(q,v) := \frac{\partial L}{\partial v}(q,v)) \tag{III.1.1}$$

is called Legendre transformation.
ii)
A Lagrangean is called singular provided that ρ_L is not surjective, that is,

$$\det((\frac{\partial^2 L}{\partial v^a \partial v^b})_{a,b=1}^m) = 0 \tag{III.1.2}$$

For singular Lagrangeans it is not possible to solve the velocities in terms of the momenta, the undelying reason being that the Lagrangean is invariant under certain symmetries.

Let $m = \dim(\mathcal{C})$ and suppose that the rank of the matrix in (III.1.2) is $m-r$ with $0 < r \le m$. By the inverse function theorem we can solve (at least locally) $m-r$ velocities for $m-r$ momenta and the remaining velocities, that is w.l.g.

$$p_A = \frac{\partial L}{\partial v^A}(q,v) \;\Rightarrow\; v^A = u^A(q^a, p_A, v^i) \tag{III.1.3}$$

where $a, b, .. = 1, ..m,\; A, B, .. = 1, .., m-r,\; i, j, .. = 1, .., r$. It follows that inserting (III.1.3) into the remaining equations $p_i = \partial L/\partial v^i$ cannot depend on the v^i any more as otherwise the rank would exceed $m-r$. We therefore obtain r equations of the form

$$p_i = [\frac{\partial L}{\partial v^i}(q,v)]_{v^A = u^A(q^a, p_A, v^j} =: \pi_i(q^a, p_A) \tag{III.1.4}$$

which show that the p_a are not independent of each other.

Definition III.1.2
i)
The functions
$$\phi_i(q^a, p_a) := p_i - \pi_i(q^a, p_A) \tag{III.1.5}$$
are called primary constraints.
ii)
The function
$$H'(q^a, p_a, v^i) := [p_a v^a - L(q^a, p_a)]_{v^a = u^a(q^a, p_A, v^i)} \tag{III.1.6}$$
is called the primary Hamiltonian corresponding to L.

Lemma III.1.1
The primary Hamiltonian is linear in v^i with coefficients ϕ_i.

Proof of Lemma III.1.1:
Differentiating the expression
$$H'(q^a, p_a, v^i) = p_A u^A(q^a, p_B, v^j) + p_i v^i - L(q^a, u^A(q^a, p_B, v^j), v^i) \tag{III.1.7}$$
by v^i we obtain
$$\begin{aligned} \frac{\partial H'(q^a, p_a, v^j)}{\partial v^i} &= [p_A - (\frac{\partial L(q^a, v^a)}{\partial v^A})_{v^A = u^A}]\frac{\partial u^A}{\partial v^i} + [p_i - (\frac{\partial L(q^a, v^a)}{\partial v^i})_{v^A = u^A}] \\ &= [p_i - \pi_i(q^a, p_A)] = \phi_i(q^a, p_a) \end{aligned} \tag{III.1.8}$$
□
We conclude that we may write
$$H'(q^a, p_a) = \tilde{H}(q^a, p_a) + v^i \phi_i(q^a, p_a) \tag{III.1.9}$$
where the new Hamiltonian $\tilde{H}$ is independent of the remaining velocities v^i.

Theorem III.1.1
The Hamiltonian equations
$$\dot{q}^a = \frac{\partial H'}{\partial p_a},\ \dot{p}_a = -\frac{\partial H'}{\partial q^a},\ 0 = \frac{\partial H'}{\partial v^i} \tag{III.1.10}$$
are equivalent to the Euler Lagrange equations
$$\dot{q}^a = v^a,\quad \frac{\partial L}{\partial q^a} = [\frac{d}{dt}\frac{\partial L}{\partial v^a}]_{v=\dot{q}} \tag{III.1.11}$$
We leave the simple proof (just use carefully the definitions) to the reader.

The phase space $\mathcal{M}$ of the constrained system is thus coordinatized by the q^a, p_a while the v^i are *Lagrange multipliers*, they do not follow any prescribed dynamical trajectory and are completely arbitrary. Our constrained phase space is equipped with the standard symplectic structure
$$0 = \{q^a, q^b\} = \{p_a, p_b\} = \{q^a, v^i\} = \{p_a, v^i\},\ \{p_a, q^b\} = \delta_a^b \tag{III.1.12}$$
and the Hamiltonian H'.

The primary constraints force the system to the submanifold of the phase space defined by $\phi_i = 0,\ i = 1,..,r$ for which we use the short hand notation $\phi = 0$. This is consistent with the dynamics if and only if that submanifold is left invariant, that is,

$$\dot{\phi}_i = \{H', \phi_i\} = \{\tilde{H}, \phi_i\} + v^j \{\phi_j, \phi_i\} \tag{III.1.13}$$

vanishes on the constraint surface $\bar{\mathcal{M}} := \mathcal{M}_{\phi=0}$ of the phase space. Now those primary constraints fall into the following three categories:
1)
$[\dot{\phi}_i]_{\phi=0} \equiv 0$ for $i = 1,..,a$ is identically satisfied for any v^i.
2i)
$[\dot{\phi}_i]_{\phi=0} \neq 0$ and $\{\phi_j, \phi_i\}_{\phi=0} = 0$ for all $j = 1,..,r$ and $i = a+1,..,b$.
2ii)
$[\dot{\phi}_i]_{\phi=0} \neq 0$ for generic v^i but the matrix $\{\phi_j, \phi_i\}_{\phi=0}$ with $j = 1,..,r;\ i = b+1,..,r$ has maximal rank $r - b$.
In case 2ii) we do not allow that the rank is smaller than $r - b$ since then we cannot find v^i in order to set $[\dot{\phi}_i]_{\phi=0} = 0$ and the theory would become inconsistent. Inconsistent theories have to be ruled out anyway.

Let us now extend the set of primary constraints by the $\phi_i := \dot{\phi}_{i-r+a}$ with $i = r+1,..,r+b-a$ and redefine r by $r \to r' := r + b - a$. Now iterate the above case analysis (notice that H' always only contains the first r constraints while $\phi = 0$ means $\phi_i = 0,\ i = 1,..,r'$) until case 2i) no longer appears ($b = a$). The iteration stops after at most $2m - r$ steps because in each step the number of (automatically functionally independent) constraints increases by at least one and $2m$ constraints constrain the phase space to a discrete set of points.

Definition III.1.3 *The constraints $\phi_i,\ i = r' - r$ are called secondary constraints. Here r' is the value of the redefined r after the last iteration step.*

It follows that at the end of the procedure we have $[\dot{\phi}_i]_{\phi=0} \equiv 0$ identically for $i = 1,..,a$ for any choice of v^i and some $0 \leq a \leq r'$ and the matrix $\{\phi_j, \phi_i\}_{\phi=0}$ with $j = 1,..,r;\ i = a+1,..,r'$ with $r' \geq r$ has maximal rank $r' - a \leq r$. Let now $v^j = v_0^j + \lambda^\mu v_\mu^j$ where $v_0^j(q^a, p_a)$ is a special solution of the inhomogeneous linear equation

$$\{\tilde{H}, \phi_i\}_{\phi=0} + v^j \{\phi_j, \phi_i\}_{\phi=0} = 0 \tag{III.1.14}$$

and $v_\mu^j(q^a, p_a),\ \mu = 1,..,r - (r' - a)$ is a basis for the general solution of the homogeneous system. We define

$$H := \tilde{H} + v_0^j \phi_j,\ \ \varphi_\mu := v_\mu^j \phi_j \tag{III.1.15}$$

Definition III.1.4
A function $f \in C^\infty(\mathcal{M})$ is called of first class provided that $\{\phi_j, f\}_{\phi=0} = 0$ for all $j = 1,..,r'$, otherwise of second class.

Lemma III.1.2
i)
The functions φ_μ, H are of first class.
ii)
The first class functions form a subalgebra of the Poisson algebra on $\mathcal{M}$.

Proof of Lemma III.1.2:
i) is clear from the construction. ii) follows by relaizing that if f, f' are first class then there exist functions f_{ij}, f'_{ij} with $i, j = 1, .., r'$ such that $\{\phi_i, f\} = f_{ij}\phi_j, \{\phi_i, f'\} = f'_{ij}\phi_j$. A short calculation then reveals that $\{\phi_i, \{f, f'\}\}_{\phi=0} = 0$.
□
Let now $H_\lambda := H + \lambda^\mu \varphi_\mu$. Since at $\phi = 0$ the finite time evolution of a function f should be indpendent of the arbitrary parameters λ^μ we require that $\{H_{\lambda_1}, .., \{H_{\lambda_N}, f\}, ..\}_{\phi=0}$ is independent of the $\lambda_1, ..\lambda_N$ for any $N = 1, 2, ...$ It is easy to see from the above lemma that this is automatically the case if f is of first class. However, since the multiple Poisson brackets contain only the first class constraints φ_μ it is actually sufficient that $\{f, \varphi_\mu\}_{\phi=0}$ for all μ.

This motivates to extend the set of first class constraints φ_μ already found to a maximal set C_μ, $\mu = 1, .., k$ with $k \geq r - (r' - a)$ and to add them to the Hamiltonian with additional lagrange multipliers. Denote the subset of the constraints ϕ_i functionally independent of the C_μ, that is, the second class constraints, by ϕ_I, $I = 1, .., r' - k$.

Definition III.1.5
i)
The set C_μ is called the set of generators of gauge transformations.
ii)
A function $f \in C^\infty(\mathcal{M})$ is called an observable provided that $\{f, C_\mu\}_{\phi=0}$ for all $\mu = 1, .., k$.
iii)
The extended Hamiltonian is defined by

$$H_\lambda = H + \lambda^\mu C_\mu \tag{III.1.16}$$

The nomenclature stems from the fact that $\{C_\mu, f\}$ can be interpreted as an infinitesimal motion generated by the flow of the Hamiltonian vector field associated with C_μ and an obeservable is invariant under this flow at least on $\bar{\mathcal{M}}$. That all first class constraints C_μ should be considered as generators of gauge transformations (so-called Dirac conjecture) and not only the φ_μ which appear in H' is motivated by the fact that only the C_μ form a closed constraint algebra (see below), however, it does not follow strictly from the formalism. That it is physically correct to proceed that way has been confirmed in countless examples though and can even be proved under some restrictions [26].

Lemma III.1.3
We have that $r' - k = 2m'$ is even and that $(\{\phi_I, \phi_J\}_{\phi=0})$ is an invertible matrix.

Proof of Lemma III.1.3:
Suppose that $(\{\phi_I, \phi_J\}_{\phi=0})$ is singular then there exist numbers $x^J \in \mathbb{C}$ such that $\{\phi_I, C_0\}_{\phi=0} = 0$ for all I where $C_0 = x^J \phi_J$. Since $\{C_\mu, C_0\}_{\phi=0}$ anyway we find $\{\phi_i, C_0\}_{\phi=0} = 0$ for all i so that C_0 is a first class constraint independent of the C_μ. This is a contradiction to the assumed maximality. It follows that $r' - k$ is even since $(\{\phi_I, \phi_J\}_{\phi=0})$ is an antisymmetric matrix.
□

Definition III.1.6
Let $c^{IJ} := ((\{\phi_K, \phi_L\})^{-1})^{IJ}$. The Dirac bracket is defined by

$$\{f, f'\}^* := \{f, f'\} + \{\phi_I, f\}c^{IJ}\{\phi_J, f'\} \tag{III.1.17}$$

Theorem III.1.2
The Dirac bracket defines a closed but degenerate two form on $\mathcal{M}$ with kernel spanned by χ_{ϕ_I} where χ_f denotes the Hamiltonian vector field of $f \in C^\infty(\mathcal{M})$ with respect to the symplectic structure determined by $\{.,.\}$.

Proof of Theorem III.1.2:
Our conventions are $i_{\chi_f}\Omega + df = 0$ and $\{f, f'\} = -i_{\chi_f} i_{\chi_{f'}}\Omega = \chi_f(f') = i_{\chi_f}(df')$ for the relation between a nondegenerate symplectic struture Ω, Hamiltonian vector field χ_f and Poisson bracket $\{.,.\}$. Also for a $p-$form $\omega = \omega_{\alpha_1..\alpha_p} dx^{\alpha_1} \wedge .. \wedge dx^{\alpha_p}$ we define exterior derivative, contraction with vector fields v and Lie derivative by

$$\begin{aligned} d\omega &= [\partial_{\alpha_1}\omega_{\alpha_2..\alpha_{p+1}}] dx^{\alpha_1} \wedge .. \wedge dx^{\alpha_{p+1}} \\ i_v\omega &= p\; v^\alpha \omega_{\alpha\alpha_1..\alpha_{p-1}} dx^{\alpha_1} \wedge .. \wedge dx^{\alpha_{p-1}} \\ \mathcal{L}_v\omega &= [i_v \cdot d + d \cdot i_v]\omega \end{aligned} \tag{III.1.18}$$

Let $\Omega = \frac{1}{2}\Omega_{\alpha\beta} dx^\alpha \wedge dx^\beta$ (here $\alpha, \beta, .. = 1, .., 2m$). Define the inverse of $\Omega_{\alpha\beta}$ by $\Omega^{\alpha\gamma}\Omega_{\gamma\beta} = \delta^\alpha_\beta$. Then it is easy to verify that $\chi^\alpha_f = \Omega^{\alpha\beta}\partial_\beta f$ and therefore $\Omega^{\alpha\beta} = -\{x^\alpha, x^\beta\}$.

We first of all verify that a nondegenerate two form is closed if and only if the associated Poisson bracket satisfies the Jacobi identity

$$\{f_{[1}, \{f_2, f_{3]}\}\} = 0 \tag{III.1.19}$$

To see this we just need to use the formula $\{f, f'\} = -\Omega^{\alpha\beta}(\partial_\alpha f)(\partial_\beta f')$ and the fact that $\delta\Omega^{-1} = -\Omega^{-1}(\delta\Omega)\Omega^{-1}$ to conclude that (III.1.19) is equivalent with $\partial_{[\alpha}\Omega_{\beta\gamma]} = 0$.

Next we verify directly from the definition for the Dirac bracket and by similar methods applied to c_{IJ} that on all of $\mathcal{M}$ the Jacobi identity

$$\{f_{[1}, \{f_2, f_{3]}\}^*\}^* = 0 \tag{III.1.20}$$

holds. Moreover

$$\{f, \phi_I\}^* = -\{\phi_I, f\}^* = 0 \tag{III.1.21}$$

for any $I = 1, .., 2m'$ and and $f \in C^\infty(\mathcal{M})$. We can therefore introduce local coordinates $x^\alpha = (x^a, x^I := \phi_I)$ with $a = 1, ..2(m - m')$, $I = 1, .., 2m'$ such that $\{x^a, x^I\} = 0$ (Darboux theorem [93], replace the ϕ_I by equivalent constraints if necessary) and define $(\Omega^*)^{\alpha\beta} := \{x^\alpha, x^\beta\}^*$. We then see that $(\Omega^*)^{aI} = (\Omega^*)^{IJ} = 0$. Define $(\Omega^*)_{ab}$ to be the inverse of $(\Omega^*)^{ab}$ and $(\Omega^*)_{aI} = (\Omega^*)_{IJ} = 0$. Then $\Omega^* = \pi^*\Omega$ where $\pi : \mathcal{M} \to \mathcal{M}'\; (x^a, x^I) \mapsto (x^a, 0)$ is the projection to the constraint manifold defined by second class constraints. That Ω^* is closed and has the anticipated kernel is now obvious.
□

Notice that for the first class constraints C_μ and the Hamiltonian H_λ we have for any $f \in C^\infty(\mathcal{M})$ that $\{C_\mu, f\}_{\phi=0} = \{C_\mu, f\}^*_{\phi=0}$ and $\{H_\lambda, f\}_{\phi=0} = \{H_\lambda, f\}^*_{\phi=0}$ (more generally this holds for any first class function). Thus, on the constraint surface the Dirac bracket defines the same equations of motion as the original bracket. Notice however that in general $\{f, f'\}_{\phi=0} \neq \{f, f'\}^*_{\phi=0}$ unless one uses a set of second class constraints ϕ_I which are themselves Darboux coordinates which is always possible to achieve but generically difficult and even unpractical. However, the Dirac bracket is easily seen to have the important property

$$(\{f_{|\mathcal{M}'}, f'_{|\mathcal{M}'}\}^*)_{|\mathcal{M}'} = (\{f, f'\}^*)_{|\mathcal{M}'} \tag{III.1.22}$$

that is, with respect a Dirac bracket we may set the second class constraints equal to zero before or after evaluating it.

Because of this and because the equations of motion and the gauge motions generated by the first class constraints are unaltered irrespective of whether we use the original Poisson bracket or the Dirac bracket we may just forget about the second class constraints for the rest of the analysis and work off the constraint surface defined by the sond class constraints while using the Dirac bracket. The reason for treating the second class constraints differently from the first class constraints is as follows:
The cleanest way to treat a constrained Hamiltonian system is to compute the full constraint surface $\bar{\mathcal{M}} = \{m \in \mathcal{M};\ \phi_i(m) = 0\ \forall\ i = 1,..,r'\}$. Since the Hamiltonian is a first class function, its Hamiltonian flow preserves the constraint surface. Since the Hamiltonian depends on arbitrary parameters, and physical observables must be independent of those, we have required that those oberservables must be independent of the Hamiltonian flow generated by the first class constraints, at least on the constraint surface. This is, however, not possible to require for the second class constraints because their Hamiltonian flow does not preserve the constraint surface. Thus, what one should do is to compute the gauge orbits $[m]$ of points m on the constraint surface (gauge invariant quantities). The manifold so obtained is called the reduced phase space $\tilde{\mathcal{M}}$ and observables are naturally functions on $\tilde{\mathcal{M}}$. The reduced phase space is automatically equipped with a symplectic structure that one obtains locally by looking for a suitable set of first class constraints and conjugate Darboux coordinates (together with a suitable choice of second class constraints as Darboux coordinates). See [51] for details. One would then quantize the reduced system.

The reason for why that is not always done is that for non-linear systems it is extremely difficult to compute $\bar{\mathcal{M}}, \tilde{\mathcal{M}}$ even classically and the reduced symplectic structure on the observables might be so complicated that it is very hard to find a representation of the associated canonical commutation relations in the quantum theory. Thus, in order to get started with the quantization Dirac has proposed to solve the constraints not before but after the quantization. Roughly speaking, we turn the constraints into operators and impose that physical states satisfy

$$\hat{C}_\mu \psi = 0 \tag{III.1.23}$$

(this equation must actually be read in a generalized sense, see section III.7). Notice that we impose this only for the first class constraints. To see why, notice that the first class constraints must satisfy a subalgebra of the Poisson algebra (we know that $\{C_\mu, C\nu\}_\phi = 0$ therefore $\{C_\mu, C_\nu\} = f_{\mu\nu}\ ^\rho C_\rho + f_{\mu\nu}\ ^I \phi_I$ for some structure functions $f_{\mu\ni}^\rho, f_{\mu\nu}^I$ and since the Poisson bracket is first class again we know that $f_{\mu\nu}^I = 0$). Therefore upon suitable operator ordering for a solution of (III.1.23) we have that

$$0 = [\hat{C}_\mu, \hat{C}_\mu]\psi = \hat{f}_{\mu\nu}\ ^\rho \hat{C}_\rho \psi \tag{III.1.24}$$

is a consistent equation. However, if we would extend (III.1.23) to second class constraints we get the contradiction

$$0 = [\hat{\phi}_I, \hat{\phi}_J]\psi \neq 0 \tag{III.1.25}$$

since the commutator is proportional to a quantization of c_{IJ} which in the worst case is a constant (in general an operator which is not constrained to vanish). Thus, one solves the second class constraints simply by restricting the argument of the wave function to the constraint surface.

Two remarks are in order:
1)
Notice that every second class constraint classically removes one degree of freedom while every first class constraint removes two since not only we delete degrees of freedom but also compute gauge orbits. However, since the number of second class constraints is always even, the reduced phase

space has always again an even number of physical degrees of freedom (otherwise it would not have a non-degenerate symplectic structure). One may then wonder how it is possible that we just impose the constraint on the state and do not compute its gauge orbit in addition. The answer is that the wave function already depends only on half of the number of kinematical degrees of freedom (configuration space). The imposition of the constraint is actually the condition that the state be gauge invariant and simultaneously the constraint operator is deleted.
2)
One may also wonder why we do not simply remove the first class constraints as well. The procedure to do this is called *gauge fixing*. Thus, we impose additional conditions $k_\mu = 0$ which ideally pick from each gauge orbit a unique representative and such that the matrix $(\{k_\mu, C_\nu\})$ is non-degenerate on the constraint surface. One may then remove the constraints C_μ by considering the system k_μ, C_μ as second class constraints and by using the associated Dirac bracket. The reason for not doing this is that it is actually very problematic: Usually functions with the required properties simply do not exist, for instance gauge orbits can be cut more than once leading to the so-called Gribov copies [200, 26]. Also, the geometric structure of the system is very much veiled and different gauge conditions may lead to different physics.

Finally, let us display a trivial example:
Consider the phase space $\mathcal{M} = T^*(\mathbb{R}^3)$ with constraints $\phi_1 = p_1,\ \phi_2 = q^2,\ \phi_3 = p_2$ where $q^a, p_a,\ a = 1, 2, 3$ are canonically conjugate configuration and momentum coordinates. It is easy to see that $C = \phi_1$ is the only first class constraint and that ϕ_2, ϕ_3 is a pair of second class constraints. For instance, functions which are independent of q^1, q^2, p_2 are first class but also the Hamiltonian $H = -(q^1)^2 + \sum_{a=1}^3[(q^a)^2 + (p_a)^2]$ and any function which is independent of q^1 is an observable but also the function $f = p_1 q^1$. The gauge motions generated by C are translations in the q^1 direction so that the value of q^1 is pure gauge. Obviously then the only second class constraint reduced phase space is $\mathcal{M}' = T^*(\mathbb{R}^2)$ while the fully reduced phase space is $\tilde{\mathcal{M}} = T^*(\mathbb{R}^1)$.

III.2 Elements of Fibre Bundle Theory

This section recalls the most important structural elements of the theory of connections on principal fibre bundles and follows closely the excellent exposition in [139] to which the reader is referred for more details. The reason for the inclusion of this section on standard material is the pivotal role that the holonomy plays in canonical quantum gravity.

Definition III.2.1
A fibre bundle over a differential manifold σ with atlas $\{U_I, \varphi_I\}$ is a quintuple (P, σ, π, F, G) consisting of a differentiable manifod P (called the total space), a differentiable manifold σ (called the base space), a differentiable surjection $\pi : P \to \sigma$, a differentiable manifold F (called the typical fibre) which is diffeomorphic to every fibre $\pi^{-1}(x)$, $x \in \sigma$ and a Lie group G (called the structure group) which acts on F on the left, $\lambda : G \times F \to F$; $(h, f) \mapsto \lambda(h, f) =: \lambda_h(f)$, $\lambda_h \circ \lambda_{h'} = \lambda_{hh'}$, $\lambda_{h^{-1}} = (\lambda_h)^{-1}$. Furthermore, for every U_I there exist diffeomorphisms $\phi_I : U_I \times F \to \pi^{-1}(U_I)$, called local trivializations, such that $\phi_{Ix} : F \to F_x := \pi^{-1}(x)$; $f \mapsto \phi_{Ix}(f) := \phi_I(x, f)$ is a diffeomorphism for every $x \in U_I$. Finally, we require that there exist maps $h_{IJ} : U_I \cap U_J \neq \emptyset \to G$, called transition functions, such that for every $x \in U_I \cap U_J \neq \emptyset$ we have $\phi_{Jx} = \phi_{Ix} \circ \lambda_{h_{IJ}(x)}$.

Conversely, given σ, F, G and the structure functions $h_{IJ}(x)$ with given left action λ on F we can reconstruct P, π, ϕ_I as follows: Define $P' = \cup_I U_I \times F$ and introduce an equivalence relation $\sim$ by saying that $(x, f) \in U_I \times F$ and $(x', f') \in U_J \times F$ for $U_I \cap U_J \neq \emptyset$ are equivalent iff $x' = x$ and $f' = \lambda_{h_{IJ}(x)}(f)$. Then $P = P'/\sim$ is the set of eqivalence classes $[(x, f)]$ with respect to this equivalence relation with bundle projection $\pi([(x, f)]) := x$ and local trivializations $\phi_I(x, f) := [(x, f)]$.

Definition III.2.2
Two bundles defined by the collections of tuples $\{(U_I, \phi_I)\}_I$ and $\{(U'_J, \phi'_J)\}_J$ respectively are said to be equivalent if the combined collection of tuples $\{(U_I, \phi_I), (U'_J, \phi'_J)\}_{I,J}$ defines a bundle again. A bundle automorphism is a diffeomorphism of P that maps whole fibres to whole fibres. Equivalently then, two bundles are equivalent if there exists a bundle automorphism which reduces to the identity on the base space. A bundle is really an equivalence class of bundles.

Notice that the transition functions satisfy the cocycle condition $h_{IJ} h_{JK} h_{KI} = 1_G$ over $U_I \cap U_J \cap U_K$ and $h_{IJ} = h_{JI}^{-1}$ over $U_I \cap U_J$. It is crucial to realize that in general h_{IJ} is not a coboundary, that is, there are in general no maps $h_I : U_I \to G$ such that $h_{IJ}(x) = h_I(x)^{-1} h_J(x)$.

Definition III.2.3
A fibre bundle is called trivial if its transition function cocycle is a coboundary.

The reason for this notation is that trivial bundles are equivalent to direct product bundles $\sigma \times F$: Given transition functions ϕ_I, it may be checked that the transition functions $\phi'_I(x, f) := \phi_I(x, \lambda_{h_I(x)^{-1}}(f))$ are actually independent of the label I and thus there is only one of them. Therefore the bundle is diffeomorphic with $\sigma \times F$.

Definition III.2.4
A local section of P is a smooth map $s_I : U_I \to P$ such that $\pi \circ s_I = id_{U_I}$. A cross section is a global section, that is, defined everywhere on σ.

Definition III.2.5
A principal G bundle is a fibre bundle where typical fibre and structure group coincide with G. On a principal fibre bundle we may define a right action $\rho:\ G\times P\to P;\rho_h(p):=\phi_I(\pi(p),h_I(p)h)$ for $p\in\pi^{-1}(U_I)$ where $h_I:\ P\to G$ is uniquely defined by $(\pi(p)=x_I(p),h_I(p)):=\phi_I^{-1}(p)$. Since G acts transitively on itself from the right, this right action is obviously transitive in every fibre and fibre preserving. $s_I^\phi(x):=\phi_I(x,1_G)$ is called the canonical local section. Conversely, given a system of local sections s_I one can construct local trivializations $\phi_I^s(x,h):=\rho_h(s_I(x))$, called canonical local trivializations.

Notice the identity $p=\rho_{h_I(p)}(s_I^\phi(\pi(p)))=\phi_I(\pi(p),h_I(p))=\phi_{I\pi(p)}(h_I(p))$ for any $p\in\pi^{-1}(U_I)$. If $U_I\cap U_J\neq\emptyset$ and $p\in\pi^{-1}(U_I\cap U_J)$ this leads to $\rho_{h_I(p)}(s_I^\phi(\pi(p)))=\rho^\phi_{h_J(p)}(s_J^\phi(\pi(p)))$. Using the fact that ρ is a right action we conclude $s_J^\phi(\pi(p))=\rho_{h_I(p)h_J(p)^{-1}}(s_I^\phi(\pi(p))$. Since the left hand side does not depend any longer on the point p in the fibre above $x=\pi(p)$ we conclude that we have a $G-$valued functions $h_{IJ}:\ U_I\cap U_J\to G,\ x\mapsto[h_J(p)^{-1}h_I(p)]_{p\in\pi^{-1}(x)}$ where the right hand side is independent of the point in the fibre. The functions h_{IJ} are actually the structure functions of P: By definition we have $\phi_{Ix}(h_I(p))=\phi_{Jx}(h_J(p))$, thus $h_I(p)=(\phi_{Ix}^{-1}\circ\phi_{Jx})(h_J(p))\lambda_{h_{IJ}(x)}(h_J(p))=h_{IJ}(x)h_J(p)$ which also shows that the left action in P reduces to left translation in the fibre coordinate.

In a principal G bundle it is easy to see, using transitivity of the right action of G, that triviality is equivalent with the existence of a global section. This is not the case for vector bundles which always have the global section $s_I(x)=\phi_I(x,0)$ but may have non-trivial transition functions.

Definition III.2.6
A vector bundle E is a fibre bundle whose typical fibre F is a vector space. The vector bundle associated with a principal G bundle P (where G is the structure group of E) under the left representation τ of G on F, denoted $E=P\times_\tau F$, is given by the set of equivalence classes $[(p,f)]=\{(\rho_h(p),\tau(h^{-1})f);\ h\in G\}$ for $(p,f)\in P\times F$. The projection is given by $\pi_E([(p,f)]):=\pi(p)$ and local trivializations are given by $\psi(x,f)=[(s_I(x),f)]$ since $[(\rho_h(s_I(x)),f)]=[(s_I(x),\tau(h)f]=[s_I(x),f')]$. Transition functions result from $u=[s_J(\pi(u)),f_J(u)]=[\rho_{h_{IJ}(\pi(u))}(s_I(\pi(u))),f_J(u))]=[(s_I(\pi(u))),\tau(h_{IJ}(\pi(u)f_J(u))]=[(s_I(\pi(u)),f_I(u)]$ and are thus gven by $\tau(\rho_{IJ}(x))$.

Conversely, given any vector bundle E we can construct a principal G bundle P such that E is associated with it by going through the above mentioned reconstruction process and by using the same structure group (with τ as the defining representation) acting on the fibre G by left translations and the same transition functions. A vector bundle is then called trivial if its associated principal fibre bundle is trivial.

Every principal fibre bundle P is naturally equipped with a vertical distribution, that is, an assignment of a subspace $V_p(P)$ of the tangent space $T_p(P)$ at each point p of P that is tangent to the fibre above $\pi(p)$. (Notice that distributions are not necessarily integrable, i.e they do not form the tangent spaces of a submanifold of P). These vertical distributions are generated by the fundamental vector fields v_Y associated with an element $Y\in$Lie(G) of the Lie algebra of G which are defined through their action on functions $f\in C^\infty(P)$:

$$(v_Y[f])(p):=(\frac{d}{dt})_{t=0}f(\rho_{\exp(tY)}(p)) \tag{III.2.1}$$

where $\exp:\ \mathrm{Lie}(G)\to G$ denotes the exponential map. The map $v:\ \mathrm{Lie}(G)\to V_p(P);\ Y\to v_Y$ is a Lie agebra homomorphism by construction.

The complement $H_p(P)$ of $V_p(P)$ in $T_p(P)$ is called the horizontal distribution and is one way to define a connection on P. More precisely

Definition III.2.7
A connection on a principal G bundle P is a distribution of horizontal subspaces $H_p(P)$ of $T_p(P)$ such that
a) $H_p(P)\oplus V_p(P)=T_p(P)$ *(i.e.* $H_p(P)\cap V_p(P)=\{0\},\ H_p(P)\cup V_p(P)=T_p(P)$*).*
b) If $v(p)=v^H(p)+v^V(p)$ *denotes the unique split of a smooth vector field into its horizontal and vertical components respectively, then the components are smooth vector fields again.*
c) $H_{\rho_h(p)}(P)=(\rho_h)_* H_p(P)$*.*

Condition c) tells us how horizontal supspaces in the same fibre are related. Here $((\rho_h)_*v)[f]=v[(\rho_h)^*f]$ denotes the push-forward of a vector field and $(\rho_h)^*f=f\circ\rho_h$ the pull-back of a function.

A different, less geometrical definition of a connection consistent with definition III.2.7 is as follows:

Definition III.2.8
A connection on a principal G bundle P is a Lie algebra valued one form ω on P which projects $T_p(P)$ into $V_p(P)$, that is
a) $\omega(v_Y)=Y$
b) $(\rho_h)^*\omega=ad_{h^{-1}}(\omega)$
c) $H_p(P)=\{v\in T_p(P);\ i_v\omega=0\}$*.*

Here $ad:\ G\times \mathrm{Lie}(G)\to Lie(G);(x,Y)\mapsto hYh^{-1}$ denotes the adjoint action of G on its own Lie algebra and i_v denotes the contraction of vector fields with forms. To see that both definitions are consistent we notice that

$$((\rho_h)^*\omega)_p(v_p)=(\omega)_{\rho_h(p)}((\rho_h)_*v_p)=(\mathrm{ad}_{h^{-1}}\omega)_p(v_p)=h^{-1}\omega_p(v_p)h \tag{III.2.2}$$

so that $v_p\in H_p(P)$ implies $(\rho_h)_*v_p\in H_{\rho_h(p)}(P)$ indeed, demonstrating that conditions b),c) of definition III.2.8 imply condition c) of definition III.2.7. Condition a) is an additional requirement fixing an otherwise free constant factor in ω.

For practical applications it is important to have a coordinate expression for ω. To that end, let us express ω in a local trivialization $p=\phi_I(x,h)$. Introducing matrix element indices $A,B,C,..$ for group elements $h=(h_{AB})$ we have

$$v_Y^\mu(p)=(\frac{\partial\phi_I^\mu(x,h)}{\partial h_{AB}}(hY)_{AB})_{\phi_I(x,h)=p} \tag{III.2.3}$$

where p^μ denotes the coordinates of p. Recalling the definition $(x_I(p)=\pi(p),h_I(p)):=\phi_I^{-1}(p)$ we claim that

$$(\omega_I(p))_{AB}=\mathrm{ad}_{h_I(p)^{-1}}(\pi^*A_I)(p)_{AB}+(h_I(p)^{-1})_{AC}dh_I(p)_{CB} \tag{III.2.4}$$

where $A_I(x)$ is a Lie(G) valued one form on U_I. Let us check that properties a), b) and c) are satisfied.

a)
We have $(\pi^*A_I)(v_Y)_p=A_I(\pi_*v_Y)_{\pi(p)}$ but $(\pi\_astv_Y)^\mu(x)=[\partial\pi^\mu(\phi_I(x,h))/\partial h_{AB}](hY)_{AB}=0$ since $\pi(\phi_I(x,h))=x$ is independent of the fibre coordinate h. On the other hand

$$\begin{aligned}&(h_I(p)^{-1}dh_I[v_Y]_p)_{AB}=h_I(p)^{-1}_{AC}[\partial h_I(p)_{CB}/\partial p^\mu][\partial\phi^\mu(x,h)/\partial h_{DE}(hY)_{DE}]_{p=\phi_I(x,h)}\\ =\ &h_I(p)^{-1}_{AD}(h_I(p)Y)_{DB}=Y_{AB}\end{aligned} \tag{III.2.5}$$

where the $h_{AB},\ A,B=1,..,\dim(G)$ could be treated as independent coordinates (although, depending on the group, this may not be the case) because of the chain rule. More precisely,

$$
\begin{aligned}
& h_I(p)^{-1}dh_I[v_Y]_p = h_I(p)^{-1}[\partial h_I(p)/\partial p^\mu][(\frac{d}{dt})_{t=0}\phi^\mu(x,he^{tY})]_{p=\phi_I(x,h)} \\
= \ & h_I(p)^{-1}(\frac{d}{dt})_{t=0}h_I(p)e^{tY} = Y
\end{aligned}
\tag{III.2.6}
$$

b)
We have $\rho_h(p)=\phi_I(\pi(p),h_I(p)h)=\phi(\pi(p),h_I(\rho_h(p))$ since ρ is fibre preserving whence $h_I(\rho_h(p))=h_I(p)h$. Since (π^*A_I) depends only on $\pi(p)$ we have $(\pi^*A_I)(\rho_h(p)) = (\pi^*A_I)(p)$. Finally, since $\rho^*d=d\rho^*$ we easily find

$$(\rho_h^*\omega)(p)=\mathrm{ad}_{h_I(p)h}(\pi^*A_I)(p)+(h_I(p)h)^{-1}dh_I(p)h=\omega(\rho_h(p))=\mathrm{ad}_{h^{-1}}(\omega(p)) \tag{III.2.7}$$

as claimed.
c)
Was already checked above.

Consider the pull-back of ω to σ by the canonical local section $s_I^\phi(x)=\phi_I(x,1_G)$. Obviously $h_I(s^\phi(x))=1_G$ whence $((s_I^\phi)^*dh_I)(x)=d1_G=0$ and $((s_I^\phi)^*\pi^*A_I)(x)=(\pi\circ s_I^\phi)^*A_I)(x)=A_I(x)$ since $\pi\circ s_I=\mathrm{id}_\sigma$ for any section. We conclude

Definition III.2.9
The so-called connection potentials

$$A_I=(s_I^\phi)^*\omega \tag{III.2.8}$$

are nothing else than the pull-back of the connection by local sections.

By its very defintion, the connection ω is globally defined therefore the above coordinate formula must be independent of the trivialization. This implies the following identity between the potentials $A_I(x)$

$$\pi^*A_I=\pi^*[\mathrm{ad}_{h_{IJ}}\pi^*A_J-dh_{IJ}h_{IJ}^{-1}] \tag{III.2.9}$$

as one can easily verify using $(\pi^*h)_{IJ}(p)=h_I(p)h_J(p)^{-1}$. We can also pull this identity back to σ and obtain

$$A_I=\mathrm{ad}_{h_{IJ}}(A_J)-dh_{IJ}h_{IJ}^{-1} \tag{III.2.10}$$

which is called the transformation behaviour of the connection potentials under a change of section (or trivialization or gauge). Since the bundle P can be reconstructed from G,σ and the transition functions $h_{IJ}(x)$ we conclude that a connection can be defined uniquely by a system of pairs consisting of connection potentials and local sections (A_I,s_I) respectvely, subject to the above transformation behaviour.

Definition III.2.10
Given a principal G bundle P over σ and a curve c in σ we define a curve $\tilde{c}$ to be the horizontal lift of c provided that
i) $\pi(\tilde{c})=c$
ii) $d\tilde{c}(t)/dt\in H_{\tilde{c}(t)}(P)$ *for any t in the domain* $[0,1]$ *of the parametrization of c.*

We now show that the lift is actually unique: We know that $\tilde{c}(t)=\phi_I(c(t),h_{cI}(t)^{-1})=\rho_{h_{cI}(t)^{-1}}(s_I^\phi(c(t))$ for some function $h_{cI}(t)$ (to be solved for) when $c(t)$ lies in the chart U_I. It follows that

$$d\tilde{c}(t)/dt=[\partial\phi_I/\partial x^a\dot{c}^a(t)+\partial\phi_I/\partial h_{AB}(\dot{h}_{cI}(t)^{-1})_{AB}](\phi_I(x,h)=\tilde{c(t)}) \tag{III.2.11}$$

That this vector is horizontal along $\tilde{c}(t)$ means that $\omega[\dot{\tilde{c}}]_{\tilde{c}(t)}=0$. Using $\omega=\mathrm{ad}_{h_I^{-1}}(\pi^*A_I)+h_I^{-1}dh_I$ we find

$$\omega[\dot{\tilde{c}}]_{\tilde{c}(t)}=h_{cI}(t)[A_{Ia}(c(t))h_{cI}(t)^{-1}\dot{c}^a(t)+\frac{d}{dt}(h_{cI}(t)^{-1})] \tag{III.2.12}$$

implying the so-called parallel transport equation (dropping the index I)

$$\dot{h}_{cI}(t)=h_{cI}(t)A_{Ia}(c(t))\dot{c}^a(t) \tag{III.2.13}$$

which is an ordinary differential equation of first order and therefore has a unique solution by the usual existence and uniqueness theorems if we provide an initial datum $\tilde{c}(0)$. The point $\tilde{c}(1)$ is called the parallel transport of $\tilde{c}(0)$. Since the point $c(1)$ in the base is already known, the essential information is contained in the group element $h_{cI}=h_{cI}(1)$ to which we will also refer to as the holonomy of A_I along c. It should be noted, however, that while $\tilde{c(1)}$ is globally defined, h_{cI} depends on the choice of the local trivialization. In fact, under a change of trivialization $A_I(x)\mapsto A_J(x)=-dh_{JI}(x)h_{JI}^{-1}(x)+\mathrm{ad}_{hJI(x)}(A_I(x))$ we obtain $h_{cJ}=h_{JI}(c(0))h_{cI}h_{JI}(c(1))^{-1}$ which maybe checked by inserting these formulas into the parallel transport equation with $x,c(1)$ replaced by $c(t)$ and relying on the uniqueness property for solutions of ordinary differential equations. It is easy to check that if c is within the domain of a chart, then an analytic formula for $h_c(A)$ is given by

$$h_c(A)=\mathcal{P}e^{\int_c A}=1+\sum_{n=1}^\infty\int_0^1 dt_n\int_0^{t_n}dt_{n-1}..\int_0^{t_2}dt_1 A(t_1)..A(t_n) \tag{III.2.14}$$

where $A(t)=A_a^j(c(t))\dot{c}^a(t)\tau_j/2$, $\tau_j/2$ is a Lie algebra basis and $\mathcal{P}$ denotes the path ordering symbol (the smallest path parameter to the left).

Definition III.2.11
Let V be a vector space and $\psi\in\bigwedge^n(P)\otimes V$ be a vector valued $n-$form on P. The covariant derivative $\nabla\psi$ of ψ is the element of $\bigwedge^{n+1}(P)\otimes V$ defined uniquely by

$$(\nabla\psi)_p[v_1,..,v_{n+1}]:=d\psi_p[v_1^H,..,v_{n+1}^H] \tag{III.2.15}$$

where $v_k\in T_p(P)$, v_k^H is its horizontal component and d is the ordinary exteriour derivative.

This definition can be applied to the connection one form where the vector space is given by $V=$Lie(G).

Definition III.2.12
The covariant derivative of the connection one-form $\omega\in\bigwedge^1(P)\otimes Lie(P)$ is called the curvature two-form $\Omega=\nabla\omega$ of ω.

The curvature inherits from ω the property

$$\rho_h^*\Omega=\mathrm{ad}_{h^{-1}}(\Omega) \tag{III.2.16}$$

To see this, notice that the property $(\rho_h)_* H_p(P) = H_{\rho_h(p)}(P)$ of the horizontal suspaces means that $(\rho_h)_* v_p^H \in H_{\rho_h(p)}(P)$ for any $v \in T_p(P)$. Since every element of $H_{\rho_h(p)}(P)$ can be obtained this way and $(\rho_h)_*$ is a bijection we conclude $[(\rho_h)_* v_p]^H = (\rho_h)_* v_p^H$. Thus

$$\begin{aligned}
& (\rho_h^*\Omega)_p(u_p, v_p) = \Omega_{\rho_h(p)}((\rho_h)_* u_p, (\rho_h)_* v_p) = d\omega_{\rho_h(p)}([(\rho_h)_* u_p]^H, [(\rho_h)_* v_p]^H) \\
= \;& d\omega_{\rho_h(p)}((\rho_h)_* u_p^H, (\rho_h)_* v_p^H) = (d\rho_h^*\omega)_p(u_p^H, v_p^H) = \mathrm{ad}_{h^{-1}}(d\omega_p)(u_p^H, v_p^H) \\
= \;& \mathrm{ad}_{h^{-1}}(\Omega_p)(u_p, v_p)
\end{aligned} \tag{III.2.17}$$

Definition III.2.13
An element $\psi \in \bigwedge^n(P) \otimes F$ is said to be of type (τ, F) (or equivariant under ρ) for some representation τ of G on F iff $\rho_h^*\psi = \tau(h)\psi$.

It follows that the curvature Ω is of type $(\mathrm{ad}, \mathrm{Lie}(G))$.

Definition III.2.14
Let $\psi \in \bigwedge^m(P) \otimes Lie(G), \xi \in \bigwedge^n(P) \otimes Lie(G)$ then

$$[\psi, \xi] := \psi \wedge \xi - (-1)^{mn} \xi \wedge \psi = \psi^j \wedge \xi^k [\tau_j, \tau_k] \in \bigwedge^{m+n}(P) \otimes Lie(G) \tag{III.2.18}$$

where τ_j is some basis of the Lie algebra of G.

Theorem III.2.1 (Cartan Structure Equation)

$$\Omega = d\omega + \omega \wedge \omega \tag{III.2.19}$$

Proof of Theorem III.2.1:
Using the split $u = u^H + u^V$ it is clear that $\omega \wedge \omega(u, v) = \omega \wedge \omega(u^V, v^V)$ because ω_p annihilates $H_p(P)$. Notice that $[\omega, \omega] = 2\omega \wedge \omega$.

Likewise we write

$$d\omega(u, v) = d\omega(u^H, v^H) + d\omega(u^H, v^V) + d\omega(u^V, v^H) + d\omega(u^V, v^V) \tag{III.2.20}$$

and use the differential geometric identity $d\omega(u, v) = u[i_v\omega] - v[i_u\omega] - i_{[u,v]}\omega$ with $(i_u\psi)(v_1, .., v_{n-1}) := \sum_{k=1}^n (-1)^{k+1} \psi(v_1, , , v_{k-1}, u, v_{k+1}, .., v_n)$ for the contraction of an $n-$form with a vector field (see, e.g. the second reference in [93]).

To evaluate these four terms in (III.2.20) we need two preliminary results:
1)
We can always find $X, Y \in \mathrm{Lie}(G)$ such that $u^V = v_X, v^V = v_Y$ are displayed as fundamental vector fields. It is easy to verify that $[u^V, v^V] = [v_X, v_Y] = v_{[X,Y]} \in V_p(P)$ is a Lie algebra homomorphism. We will exploit that $\omega(u^V) = X$ etc. is a constant.
2)
By definition of the Lie bracket of vector fields $[u^V, v^H] = (d/dt)_{t=0}[\rho_{h^{u^V}(t)}]_* v^V \in H_p(P)$ since the push-forward by the right action preserves horizontal vector fields ($h^{u^V}(t)$ denotes the integral curve of u^V). We will exploit that $\omega(w^H) = 0$ for any horizontal vector field w^H.

Using these two properties it is immediate that $d\omega(u^H, v^V) = d\omega(u^V, v^H) = 0$ and that $d\omega(u^V, v^V) = -\omega([v_X, v_Y]) = -[X, Y]$. On the other hand

$$\omega \wedge \omega(u^V, v^V) = i_{v^V} i_{u^V} \omega \wedge \omega i_{v^V}[\omega(u_v)\omega - \omega\omega(u_v)] = [\omega(v_X), \omega(v_Y)] = +[X, Y]$$

Therefore we are left with

$$[d\omega + \omega \wedge \omega](u, v) = d\omega(u^H, v^H) = \Omega(u, v) \tag{III.2.21}$$

□

Corollary III.2.1 (Bianchi Identity)

$$\nabla \Omega = 0 \tag{III.2.22}$$

To prove this, use the Cartan structure equation to infer $d\omega = \omega \wedge \omega - \omega \wedge d\omega = [\omega \wedge \Omega]$ and use $\omega(u^H) = 0$ again.

Definition III.2.15
The local field strength $F_I := 2s_I^*\Omega = 2[dA_I + A_I \wedge A_I$ is twice the pull-back by local sections of the curvature two-form.

Using the transformation behaviour of the connection potential under a change of trivialization it is easy to verify the the corresponding change of the field strength is given by

$$F_J(x) = \mathrm{ad}_{g_{JI}(x)}(F_I(x)) \tag{III.2.23}$$

whence traces of polynomials in the field strength, used in classical action principles of gauge field theories are globally defined (gauge invariant).

Definition III.2.16
Let $E = P \times_\tau F$ be a vector bundle associated to P, c a curve in σ and $\tilde{c}$ its horizontal lift which we display as above as $\tilde{c}(t) = \rho_{h_{cI}(t)^{-1}}(s_I^\phi(c(t))$. A local section of E is then given by $S_I(x) = [(s_I^\phi(x), f_I(x))]$ where $f_I(x)$ is called the fibre section, whence

$$S_I(c(t)) = [(s_I^\phi(c(t)), f_I(c(t))] = [\rho_{h_{cI}(t)^{-1}}(s_I^\phi(c(t)), \tau(h_{cI}(t))f_I(c(t))] = [\tilde{c}(t), \tau(h_{cI}(t))f_I(c(t))] \tag{III.2.24}$$

The covariant differential of S_I along $v := \dot{c}(0)$ at $x = c(0)$ is defined by

$$(\nabla_v S_I)_x := [\tilde{c}(0), (\frac{d}{dt})_{t=0}\tau(h_{cI}(t))f_I(c(t))] \tag{III.2.25}$$

It is easy to see, using the equivalence relation in the definition of E and the definition of the horizontal lift that (III.2.25) is actually independent of the initial datum for $\tilde{c}$ or, equivalently, the group element h_0 in $\tilde{c}(0) = \rho_{h_0}(s_I(x), 1_G)$. Notice that multiplication of sections by scalar functions is defined by $f(x)S_I(x) = [(s_I^\phi(x), f(x)f_I(x)]$ so that the covariant differential ∇ satisfies the usual axioms for a covariant differential (Leibniz rule).

As usual, one is interested for practical calculations in coordinate expressions. To that end, consider a constant basis e_α in F and consider the special sections $S_{I\alpha}(x) := [(s_I^\phi(x), e_\alpha)]$. From the differential equation for the holonomy (III.2.13) with initial condition $h_{cI}(0) = 1_G$ we conclude

$$\begin{aligned}(\nabla_v S_{I\alpha})(x) &= [(s_I^\phi(x), (\frac{\partial\tau(h)}{\partial h_{AB}}]_{h=1_G}(A_{Ia}(x))_{AB} v^a e^\alpha)] \\ &= v^a A^j_{Ia}(x)[(s_I^\phi(x), (\frac{d\tau(\exp(t\tau_j))}{dt})_{t=0} e_\alpha)] = v^a A^j_{Ia}(x)\tau^\tau_j S_{I\alpha}(x)\end{aligned} \tag{III.2.26}$$

where we have abbreviated by $\tau^\tau_j = (\frac{d\tau(\exp(t\tau_j))}{dt})_{t=0}$ a basis of Lie(G) in the representation τ and have expanded $A_I = A^j_I \tau_j$ correspondingly. Using the Leibniz rule and the fact that a general section may be written as $S_I(x) = f^\alpha_I(x) S_{I\alpha}(x)$ we find

$$\nabla_v S_I = i_v[df^\alpha_I S_{I\alpha} + f^\alpha_I A^j_I \tau^\tau_j S_{I\alpha}] \tag{III.2.27}$$

This expression becomes especially familiar if we use the standard basis $(e_\alpha)^\beta = \delta^\beta_\alpha$ whence $f^\alpha_I(Me_\alpha) = M^\beta_\alpha f^\alpha_I e_\beta = (Mf_I)^\alpha e_\alpha$ for any matrix M so that

$$\nabla_v S_I = i_v[df_I + A^j_I \tau^\tau_j f_I]^\alpha S_{I\alpha} =: [i_v(\nabla f_I)^\alpha] S_{I\alpha} \tag{III.2.28}$$

We now require that $S_I = S$ is actually globally defined which will require a certain transformation behaviour of $f_I(x)$ under a change of section. We have $p = \rho_{h_I(p)}(s^\phi_I(x)) = \rho_{h_J(p)}(s^\phi_J(x))$ so that $s^\phi_J(x) = \rho_{h_{IJ}(x)}(s^\phi_I(x))$, thus $S_J(x) = [(s^\phi_I(x), \tau(h_{IJ}(x)) f_J(x))] = S_I(x)$ requires that the fibre section transforms as

$$f_J(x) = \tau(h_{JI}(x)) f_I(x) \tag{III.2.29}$$

This leads to the following covariant transformation property of its covariant derivative ($c(0) = x, \dot{c}(0) = v$):

$$\begin{aligned}(\nabla_v f_J)(x) &= i_v(df_J)_x + (\frac{d}{dt})_{t=0}\tau(h_{cJ}(t)) f_J(x) \\ &= \tau(h_{JI}(x))[i_v(df_I)_x + \tau(h_{JI}(x))^{-1}[i_v d\tau(h_{JI})](x) f_I(x) \\ &\quad + \tau(h_{JI}(x))^{-1}(\frac{d}{dt})_{t=0}\tau(h_{JI}(x)h_{cI}(t)h_{JI}(c(t))^{-1})\tau(h_{JI}(x))f_I(x)] \\ &= \tau(h_{JI}(x))[i_v(df_I)_x + \tau(h_{JI}(x))^{-1}[i_v d\tau(h_{JI})](x) f_I(x) \\ &\quad + (\frac{d}{dt})_{t=0}\tau(h_{cI}(t)h_{JI}(c(t))^{-1}) f_I(x) \\ &\quad + (\frac{d}{dt})_{t=0}\tau(h_{JI}(c(t))^{-1})\tau(h_{JI}(x)) f_I(x)] \\ &= \tau(h_{JI}(x))[(\nabla_v f_I)(x) + \{\tau(h_{JI}(x))^{-1}[i_v d\tau(h_{JI})](x) f_I(x) \\ &\quad + [i_v d\tau(h_{JI})^{-1}](x)\tau(h_{JI}(x))\} f_I(x)] \\ &= \tau(h_{JI}(x))(\nabla_v f_I)(x)\end{aligned} \tag{III.2.30}$$

which implies that the cross section S has a globally defined covariant differential.

Definition III.2.17 *A cross section S in $E = P \times \tau F$ is said to be parallel transported along a curve c in σ iff $(\nabla_{\dot{c}(t)} S)(c(t)) = 0$ for all $t \in [0,1]$.*

Notice that we may consider the covariant differential as a map $\nabla : \mathcal{S}(E) \to \mathcal{S}(E) \otimes \bigwedge^1(\sigma)$ where $\mathcal{S}(E)$ denotes the space of sections of E. We extend this definition to $\nabla : \mathcal{S}(E) \otimes \bigwedge^n(\sigma) \to \mathcal{S}(E) \otimes \bigwedge^{n+1}(\sigma)$ through the "Leibniz rule"

$$\nabla(S \otimes \psi) := (\nabla S) \wedge \psi + S \otimes d\psi \tag{III.2.31}$$

This way we can rediscover the field strength through the square of the covariant differential:

$$\begin{aligned} \nabla^2 S &= \nabla^2 S_\alpha \otimes f^\alpha = \nabla[\nabla S_\alpha \otimes f^\alpha + S_\alpha \otimes df^\alpha] \\ &= \nabla S_\alpha \otimes [df^\alpha + A^\alpha_\beta f^\beta] = S_\alpha \otimes \{A^\alpha_\gamma \wedge [df^\gamma + A^\gamma_\beta f^\beta] + d(A^\alpha_\beta f^\beta)\} \\ &= = S_\alpha \otimes [dA^\alpha_\beta + A^\alpha_\gamma \wedge A^\gamma_\beta] f^\beta = \frac{1}{2} S_\alpha \otimes F^\alpha_\beta f^\beta \end{aligned} \tag{III.2.32}$$

III.3 Tools from General Topology

We collect and prove here some important results from general topology needed in the main text. For more details, see e.g. [140].

Definition III.3.1
I)
i)
Let X be a set and $\mathcal{U}$ a collection of subsets of X. We call X a topological space provided that
1) $\emptyset, X \in \mathcal{U}$
2) $\mathcal{U}$ is closed under finite intersections: $U_1, .., U_N \in \mathcal{U}, N \in \mathbb{N} \Rightarrow \bigcap_{k=1}^N U_k \in \mathcal{U}$
3) $\mathcal{U}$ is closed under arbitrary (possibly uncountably infinite) unions: $U_\alpha \in \mathcal{U}, \alpha \in A \Rightarrow \bigcup_{\alpha \in A} U_\alpha \in \mathcal{U}$
The sets $U \in \mathcal{U}$ are called open, their complements $X - U$ closed in X. If $x \in X$ is a point and U an open set containing it then U is called a neighbourhood of x in X. A topology $\mathcal{U}$ is called stronger (finer) then a topology $\mathcal{U}'$ which then is weaker (coarser) if $\mathcal{U}' \subset \mathcal{U}$.
ii)
Let $(X, \mathcal{U}), (Y, \mathcal{V})$ be topological spaces such that $Y \subset X$. The relative or subspace topology $\mathcal{U}_Y$ induced on Y is given by defining the sets $U \cap Y$; $U \in \mathcal{U}$ to be open. We say that we have a topological inclusion, denoted $Y \hookrightarrow X$, provided that the intrinsic topology is stronger than the relative one, that is, $\mathcal{U}_Y \subset \mathcal{V}$.
II)
i)
A function $f : X \to Y$ between topological spaces X, Y is said to be continuous provided that the preimage $f^{-1}(V)$ of any set $V \subset Y$ that is open in Y is open in X. (The preimage is defined by $f^{-1}(V) = \{x \in X;\ f(x) \in V\}$ and despite the notation does not require f to be either an injection or a surjection). One easily shows that f is continuous if it is continuous at each point $x \in X$. Here f is continuous at $x \in X$ if for any open neighbourhood V of $y = f(x)$ there exists an open neighbourhood U of x such that $f(x') \in V$ for all $x' \in U$ (i.e. $f(U) \subset V$).
ii)
If f is a continuous bijection and also f^{-1} is continuous then f is called a homeomorphism or a topological isomorphism.

We see that a topology on a set X is simply defined by saying which sets are open, or equivalently, which functions are continuous. The importance of homeomorphisms f for topology is that not only the spaces X, Y can be identified set theoretically but also topologically, that is, open sets can be identified with each other.

In order to get more topological spaces with more structure one must add separation and compactness properties. The one we need here is the following.

Definition III.3.2
i)
A topological space X is said to be Hausdorff iff for any two of its points $x \neq y$ there exist neighbourhoods U, V of x, y respectively which are disjoint.
ii)
A topological space X is called compact if every open cover $\mathcal{V}$ of X (a collection of open sets of X whose union is all of X) has a finite subcover.

Definition III.3.3
i)
A net (x^α) in a topological space X is a map $\alpha \to x^\alpha$ from a partially ordered and directed index set A (relation $\geq$) to X.
ii)
A net (x^α) converges to x, denoted $\lim_\alpha x^\alpha = x$ if for every open neighbourhood $U \subset X$ of x there exists $\alpha(U) \in A$ such that $x^\alpha \in U$ for every $\alpha \geq \alpha(U)$ (one says that (x^α) is eventually in U).
iii)
A subnet $(x^{\alpha(\beta)})$ of a net (x^α) is defined through a map $B \to A;\ \beta \mapsto \alpha(\beta)$ between partially ordered and directed index sets such that for any $\alpha_0 \in A$ there exists $\beta(\alpha_0) \in B$ with $\alpha(\beta) \geq \alpha_0$ for any $\beta \geq \beta(\alpha_0)$ (one says that B is cofinal for A).
iv)
A net (x^α) in a topological space X is called universal if for any subset $Y \in X$ the net (x^α) is eventually either only in Y or only in $X - Y$.

Notice that for a subnet there is no relation between the index sets A, B except that $\alpha(B) \subset A$ so that in particular the subnet of a sequence ($A = \mathbb{N}$) may not be a sequence any longer. The notions of closedness, continuity and compactness can be formulated in terms of nets. The fact that one uses nets instead of sequences is that lemma III.3.1 is no longer true when $A = \mathbb{N}$ unless we are dealing with metric spaces.

Lemma III.3.1
i) A subset Y of a toplogical space X is closed if for every convergent net (x^α) in X with $x^\alpha \in Y\ \forall\alpha$ the limit actually lies in Y.
ii)
A function $f:\ X \to Y$ between topological spaces is continuous if for every convergent net (x^α) in X, the net $(f(x^\alpha))$ is convergent in Y.
iii)
A topological space X is compact if every net has a convergent subnet. The limit point of the convergent subnet is called a cluster (accumulation) point of the original net.

The proof is standard and will be omitted. One easily sees that if a net converges (a function is continuous) in a certain topology, then it does so in any weaker (stronger) topology. In our applications direct products of topological spaces are of fundamental importance.

Definition III.3.4
The Tychonov topology on the direct product $X_\infty = \prod_{l\in\mathcal{L}} X_l$ of topological spaces X_l, $\mathcal{L}$ any index set, is the weakest topology such that all the projections

$$p_l:\ X_\infty \to X_l;\ (x_{l'})_{l'\in\mathcal{L}} \mapsto x_l \tag{III.3.1}$$

are continuous, that is, a net $x^\alpha = (x^\alpha_l)_{l\in\mathcal{L}}$ converges to $x = (x_l)_{l\in\mathcal{L}}$ iff $x^\alpha_l \to x_l$ for every $l \in \mathcal{L}$ pointwise (not necessarily uniformly) in $\mathcal{L}$. Equivalently, the sets $p_l^{-1}(U_l) = [\prod_{l'\neq l} X_{l'}] \times U_l$ are defined to be open and form a base for the topology of X_∞ (any open set can be obtained from those by finite intersections and arbitrary unions).

The definition of this topology is motivated by the following theorem.

Theorem III.3.1 (Tychonov)
Let $\mathcal{L}$ be an index set of arbitrary cardinality and suppose that for each $l \in \mathcal{L}$ a compact topological space X_l is given. Then the direct product space $X_\infty = \prod_{l\in\mathcal{L}} X_l$ is a compact topological space in the Tychonov topology.

We will give an elegant proof of the Tychonov theorem using the notion of a universal net.

Lemma III.3.2
i)
A universal net has at most one cluster point to which it then converges.
ii)
For any map $f : X \to Y$ between topological spaces the net $f(x^\alpha)$ in Y is universal whenever (x^α) is universal in x with no restrictions on f.
iii)
Any net has a universal subnet.

Proof of Lemma III.3.2:
i)
Suppose that x is a cluster point of a universal net (x^α) and that the subnet $x^{\alpha(\beta)}$ converges to it. Thus for any neighbourhood U of x the subnet is eventually in U, i.e. there exists $\beta(U)$ such that $x^{\alpha(\beta)} \in U$ for any $\beta \geq \beta(U)$. Since (x^α) is universal it must be eventually either in U or $X - U$. Suppose there was α_0 such that $x^\alpha \in X - U$ for any $\alpha \geq \alpha_0$. By definition of a subnet we find $\beta(\alpha_0)$ such that $\alpha(\beta) \geq \alpha_0$ for any $\beta \geq \beta(\alpha_0)$. Without loss of generality we may choose $\beta(\alpha_0) \geq \beta(U)$. But then we know already that the $x^{\alpha(\beta)}$, $\beta \geq \beta(\alpha_0)$ are in U which is a contradiction. Thus x^α is eventually in U. Since U was an arbitray neighbourhood of x, it follows that (x^α) actually converges to x.
ii)
Obviously $f(x^\alpha)$ is eventually in $f(X)$ so we must show that for any $V \subset f(X)$ we have $f(x^\alpha)$ eventually in V or $f(X) - V$. Let $U = f^{-1}(V)$ be the preimage of V, then $f(X - U) = f(X) - V$. Since (x^α) is eventually in U or $X - U$, the claim follows.
iii)
The proof can be found in exercise 2J d) together with theorem 2.5 in [141].
□

Corollary III.3.1
A topological space X is compact iff every universal net converges.

Proof of Corollary III.3.1:
$\Rightarrow$:
Take any universal net (x^α). Since X is compact it has a cluster point to which it actually converges by lemma III.3.1 i).
$\Leftarrow$:
Take any net (x^α). Then by lemma III.3.1iii) it has a universal subnet $x^{\alpha(\beta)}$ which converges by assumtion. Thus, X is compact.
□

Proof of Theorem III.3.1:
Let $(x^\alpha) = (x_l^\alpha)_{l\in\mathcal{L}}$ be any universal net in $X_\infty = \prod_{l\in\mathcal{L}} X_l$. By lemma III.3.1ii) the net $p_l((x^\alpha)) = (x_l^\alpha)$ is universal in X_l. Since X_l is compact, it converges to some x_l. Define $x := (x_l)_{l\in\mathcal{L}}$. By defintion of the Tychonov topology, $x^\alpha \to x$ iff $x_l^\alpha \to x_l$ for any $l \in \mathcal{L}$ whence (x^α) converges.
□
This proof of the Tychonov theorem is shorter than the usual one in terms of the (in)finite intersection property and technically clearer.

Definition III.3.5
Let Y be a subset of a topological space X. The subset topology induced by X on Y is defined through the collection of open sets $\mathcal{V} := \{U \cap Y;\ U \in \mathcal{U}\}$ where $\mathcal{U}$ defines the topology of X.

Lemma III.3.3
A closed subset Y of a compact topological space X is compact in the subspace topology.

Proof of Lemma III.3.3:
Let $\mathcal{V}$ be any open cover for Y. Since Y is closed in X, $X - Y$ is open in X whence $\mathcal{U} = \mathcal{V} \cup \{X - Y\}$ is an open cover for X. Since X is compact, it has a finite open subcover $\{U_k\}_{k=1}^N \cup \{X - Y\}$ for some $N < \infty$ where U_k is open in X. By defintion of the supspace topology, $U_k \cap Y$ is open in Y so that $\{U_k \cap Y\}_{k=1}^N$ is a finite open subcover of $\mathcal{V}$.
$\square$

In our discussion of the gauge orbit of connections we will deal with the quotient of connections by the set of gauge transformations which is a topological space again. The resulting quotient space carries a natural topology, the quotient topology.

Definition III.3.6
i)
Let X, Y be topological spaces and $p:\ X \to Y$ a surjection. The map p is said to be a quotient map provided that $V \subset Y$ is open in Y if and only if $p^{-1}(V)$ is open in X.
ii)
If X is a topological space, Y a set and $p:\ X \to Y$ a surjection then there exists a unique topology on Y with respect to which p is a quotient map.
iii)
Let X be a topological space and let $[X]$ be a partition of X (i.e. a collection of mutually disjoint subsets of X whose union is X). Denote by $[x]$, $x \in X$ the subset of X in that partition of X which contains x. Equip $[X]$ with the quotient topology induced by the map $[]:\ X \to [X];\ x \mapsto [x]$. Then $[X]$ is called the quotient space of X.

Notice that the requirement for p to be a quotient map is stronger than that it be continuous which would only require that $p^{-1}(V)$ is open in X whenever V is open in Y (but not vice versa). Clearly in ii) we define the topology on the set Y to be those subsets V for which the preimage $p^{-1}(V)$ is open in X and it is an elementary exercise in the theory of mappings of sets to verify that the collection of subsets of Y so defined satisfies the axioms of a topology of definition III.3.1.

Quotient spaces naturally arise if we have a group action $\lambda:\ G \times X \to X;\ (g, x) \to \lambda_g(x) := \lambda(g, x)$ on a topological space xX and define $[x] := \{\lambda_g(x);\ g \in G\}$ to be the orbit of x. The orbits clearly define a partition of X.

Lemma III.3.4
Let X be a compact topological space, Y a set and $p:\ X \to Y$ a surjection. Then Y is compact in the quotient topology.

Proof of Lemma III.3.4:
First of all, consider any subsets V_1, V_2 of Y.
On the one hand suppose $x \in p^{-1}(V_1) \cap p^{-1}(V_2)$. Then there exist $y_1 \in V_1, y_2 \in V_2$ such that $y_1 = p(x) = y_2$, that is, $y_1 = y_2 \in V_1 \cap V_2$ so that actually $x \in p^{-1}(V_1 \cap V_2)$. We conclude $p^{-1}(V_1) \cap p^{-1}(V_2) \subset p^{-1}(V_1 \cap V_2)$.

On the other hand, let $x \in p^{-1}(V_1 \cap V_2)$, then there exists $y \in V_1 \cap V_2$ such that $x \in p^{-1}(y)$. Since $y \in V_1 \cap V_2$ we have $p^{-1}(y) \in p^{-1}(V_1)$ and $p^{-1}(y) \in p^{-1}(V_2)$, thus $x \in p^{-1}(V_1) \cap P^{-1}(V_2)$. We conclude $p^{-1}(V_1 \cap V_2) \subset p^{-1}(V_1) \cap p^{-1}(V_2)$.
Thus, altogether $p^{-1}(V_1) \cap p^{-1}(V_2) = p^{-1}(V_1 \cap V_2)$ and $p^{-1}(V_1) \cup p^{-1}(V_2) = p^{-1}(V_1 \cup V_2)$ by taking complements.

Next, let $\mathcal{V}$ be an open cover of Y. Then, by definition of the quotient topology, $p^{-1}(V)$ is open in X and $\mathcal{U} := \{p^{-1}(V);\ V \in \mathcal{V}\}$ covers X because $\bigcup_{U \in \mathcal{U}} U = \bigcup_{V \in \mathcal{V}} p^{-1}(V) = p^{-1}(\bigcup_{V \in \mathcal{V}} V) = p^{-1}(Y) = X$ since p is a surjection and $\mathcal{V}$ covers Y. We conclude that $\mathcal{U}$ is an open cover of X.

Since X is compact, we find a finite, open subcover $\{p^{-1}(V_k)\}_{k=1}^N$ of X so that $X = \bigcup_{k=1}^N p^{-1}(V_k) = p^{-1}(\bigcup_{k=1}^N V_k) = p^{-1}(Y)$ whence $Y = \bigcup_{k=1}^N V_k$, that is, $\{V_k\}_{k=1}^N$ is a finite open subcover of $\mathcal{V}$ and Y is compact.
$\square$

Lemma III.3.5
Let X be a Hausdorff space and $\lambda :\ G \times X \to X$ a continuous group action on X (i.e., λ_g defined by $\lambda_g(x) := \lambda(g, x)$ is continuous for any $g \in G$). Then the quotient space $X/G := \{[x];\ x \in X\}$ defined by the orbits $[x] = \{\lambda_g(x);\ g \in G\}$ is Hausdorff in the quotient topology.

Proof of Lemma III.3.5:
Let $[x] \neq [x']$ then certainly $x \neq x'$ since orbits are disjoint. Since X is Hausdorff we find disjoint open neighbourhoods U, U' of x, x' respectively. We want to show that U, U' can be chosen in such a way that

$$[U] := \{[y];\ y \in U\},\ [U'] := \{[y'];\ y' \in U'\} \tag{III.3.2}$$

are disjoint. First of all we notice that (p the projection map)

$$\begin{aligned} p^{-1}([U]) &= \bigcup_{y \in U} p^{-1}([y]) = \{\lambda(g, y);\ y \in U, g \in G\} = \bigcup_{g \in H} \lambda_g(U) = \bigcup_{g \in H} \lambda_{g^{-1}}(U) \\ &= \bigcup_{g \in H} (\lambda_g)^{-1}(U) \end{aligned} \tag{III.3.3}$$

where we have made use of $\lambda_{g^{-1}} = (\lambda_g)^{-1}$. Since U is open in X and λ_g is continuous by assumption, we have that $\lambda_g^{-1}(U)$ is open in X. Since arbitrary unions of open sets are open it follows that $p^{-1}([U])$ is open in X, thus by the definition of the quotient topology we have $[U], [U']$ open in X/G. Next, obviously $[x] \in [U], [x'] \in [U']$ whence $[U], [U']$ are open neighbourhoods of $[x], [x']$ in X/G respectively.

Let us now choose V, V' to be open, disjoint neighbourhoods of the orbits $p^{-1}([x]) = \lambda_G(x), p^{-1}([x'])$ respectively. (This is certainly possible as otherwise there exists $g \in G$ such that $\lambda_g(x), x'$ have no disjoint neighbourhoods which is impossible because $\lambda_g(x) \neq x'$ (otherwise $[x] = [x']$) and X is Hausdorff). We claim that we can choose U, U' in such a way that $p^{-1}[U] := \bigcup_{g \in G} \lambda_g(U) \subset V$ and $p^{-1}[U'] := \bigcup_{g \in G} \lambda_g(U') \subset V'$.

Suppose that were not the case. Then for any neighbourhood U of x we find $z \in U$ and $g_0 \in G$ such that $\lambda_{g_0}(z) \notin V$. Since by construction of V we have that V is a common open neighbourhood of any $\lambda_g(x),\ g \in G$ we have in particular $y := \lambda_{g_0}(x) \in V$. It follows that we have found an open neighbourhood V of $y = \lambda_{g_0}(x)$ such that for any open neighbourhood U of x there exists $z \in U$ with $\lambda_{g_0}(z) \notin V$. This means that the map λ_{g_0} is not continuous at x in contradiction to our assumption that λ_g is everywhere continuous for any $g \in G$.

Therefore $p^{-1}([U]) \cap p^{-1}([U']) = p^{-1}([U] \cap [U']) = \emptyset$ whence $[U] \cap [U'] = \emptyset$, thus X/G is Hausdorff.
$\square$

Theorem III.3.2
Let X, Y be topological spaces and let G be a group acting (not necessarily continuously) on them via λ, λ' respectively. If $f: X \to Y$ is a homeomorphism with respect to which the actions λ, λ' are equivariant then f extends as a homeomorphism to the quotient spaces $X/G, Y/G$ in their respective quotient topologies.

Proof of Theorem III.3.2:
Equivariance means that $f \circ \lambda_g = \lambda'_g \circ f$ for all $g \in G$ and since f is a bijection, equivariance implies also $\lambda_g \circ f^{-1} = f^{-1} \circ \lambda'_g$ Consider the corresponding quotient maps

$$p: X \to X/G;\ x \mapsto [x]_\lambda = \{\lambda_g(x);\ g \in G\} \text{ and } p': Y \to Y/G;\ x \mapsto [x]_{\lambda'} = \{\lambda'_g(y);\ g \in G\} \tag{III.3.4}$$

Then due to equivariance

$$f([x]_\lambda) = \{f(\lambda_g(x));\ g \in G\} = \{\lambda'_g(f(x));\ g \in G\} = [f(x)]_{\lambda'} \tag{III.3.5}$$

and similarly $f^{-1}([y]_{\lambda'}) = [f^{-1}(y)]_\lambda$ so that f extends to a bijection between the corresponding equivalence classes.

Next we notice that $p^{-1}([x]_\lambda) = \{\lambda_g(x);\ g \in G\}$ whence by (III.3.5) we have $f(p^{-1}([x]_\lambda)) = (p')^{-1}([f(x)]_{\lambda'})$ for all $[x]_\lambda \in X/G$. This shows that equivariance also implies

$$f \circ p^{-1} = (p')^{-1} \circ f \ \Rightarrow\ f^{-1} \circ (p')^{-1} = p^{-1} \circ f^{-1} \tag{III.3.6}$$

Let then B be open in Y/G, thus $(p')^{-1}(B)$ open in Y by definition of the quotient topology in Y/G, thus $(f^{-1} \circ (p')^{-1})(B) = (p^{-1} \circ f^{-1})(B)$ open in X since f is continuous, thus $f^{-1}(B)$ open in X/G by definition of the quotient topology in X/G. Likewise we see that A open in X/G implies $f(A)$ open in Y/G since f^{-1} is continuous. It follows that f, f^{-1} are continuous as maps between $X/G, Y/G$.
□

III.4 Elementary Introduction to Gel'fand Theory for Abelean C^* Algebras

There are many good mathematical textbooks on operator algebra – and abstract C^*-algebra theory, see e.g. [142, 39]. The textbooks [143] are more geared towards applications in mathematical physics. For a pedagogical introduction with elegant proofs the beautiful review [144] is recommended.

Definition III.4.1
i)
An algebra $\mathcal{A}$ is a vector space (taken over $\mathbb{C}$) together with a multiplication map $\mathcal{A}\times\mathcal{A} \to \mathcal{A}$; $(a, a') \mapsto aa'$ which is associative, $(ab)c = a(bc)$, and distributive, $b(za + z'a') = zba + z'ba'$, $(za + z'a')b = zab + z'a'b$ for all $a, a', b \in \mathcal{A}$, $z, z' \in \mathbb{C}$.
ii)
An algebra $\mathcal{A}$ is called Abelean if all elements commute with each other and unital if it has a (necessarily unique) unit element 1 satisfying $1a = a1 = a$ for all $a \in \mathcal{A}$.
iii)
A vector subspace $\mathcal{B}$ of $\mathcal{A}$ is called a subalgebra if it is closed under multiplication. A subalgebra $\mathcal{I}$ is called a left (right) ideal if $ab \in \mathcal{I}$ ($ba \in \mathcal{I}$) for all $a \in \mathcal{A}$, $b \in \mathcal{I}$ and a two-sided ideal (or simply ideal) if it is simultaneously a left – and right ideal. An ideal of either kind is called maximal if there is no other ideal containing it except for $\mathcal{A}$ itself.
iv)
An involution on an algebra $\mathcal{A}$ is a map $* : \mathcal{A} \to \mathcal{A}; a \mapsto a^*$ satisfying
1) $(za + z'b)^* = \bar{z}a^ast + \bar{z}'b^*$ (conjugate linear),
2) $(ab)^* = b^*a^*$ (reverses order) and
3) $(a^*)^* = a$ (squares to the identity)
for all $a, b \in \mathcal{A}$, $z, z' \in \mathbb{C}$. An algebra with involution is called an $^*-$algebra.
v)
A homomorphism ($^*-$homomorphism) is a linear map $\Phi : \mathcal{A} \to \mathcal{B}$ between algebras ($^*-$algebras) that preserves the multiplicative (and involutive) structure, that is, $\phi(ab) = \phi(a)\phi(b)$ (and $\phi(a^*) = (\phi(a))^*$).
vi)
A normed algebra $\mathcal{A}$ is equipped with a norm $||.|| : \mathcal{A} \to \mathbb{R}^+$ (that is $||a + b|| \le ||a|| + ||b||$, $||za|| = |z|\ ||a||$, $||a|| = 0 \Leftrightarrow a = 0$, if the last property is drooped, then $||.||$ is only a seminorm) whose compatibility with the mutiplicative structure is contained in the submultiplicativity requirement $||ab|| \le ||a||\ ||b||$ for all $a, b \in \mathcal{A}$. If $\mathcal{A}$ has an involution we require $||a^*|| = ||a||$ and $\mathcal{A}$ is called a normed $^*-$algebra. If $\mathcal{A}$ is unital we require $||1|| = 1$ (this is just a choice of normalization).
vii)
A norm induces a metric $d(a, b) = ||a - b||$ and if the algebra $\mathcal{A}$ is complete (every Cauchy sequence converges) then it is called a Banach algebra.
viii)
A C^*-algebra $\mathcal{A}$ is a Banach algebra with involution with the following compatibility condition between the involutive and metrical structure

$$||a^*a|| = ||a||^2 \tag{III.4.1}$$

The innocent looking condition (III.4.1) determines much of the structure of C^*-algebras. If a C^*-algebra is not unital one can always embed it isometrically into a larger unital C^*-algebra (see e.g. [144]). While this does not remove all problems with C^*-algebras without identity in our applications only unital C^*-algebras will appear and this is what we will assume form now on. If $\mathcal{I}$ is a two-sided ideal in an algebra $\mathcal{A}$ we can form the quotient algebra $\mathcal{A}/\mathcal{I}$ which consists of the equivalence classes $[a] := \{a + b;\ b \in \mathcal{I}\}$ for any $a \in \mathcal{A}$ in which the rules for addition, multiplication and scalar multiplication are given by $[a] + [a'] = [a + a'],\ [a][a'] = [aa'],\ [za] = z[a]$ and it is easy to see that the condition that $\mathcal{I}$ is an ideal is just sufficient for making these rules independent of the representative. Finally, if we think of $\mathcal{A}$ as an algebra of operators on a Hilbert space and $||.||$ is the uniform operator norm then we see that we are dealing with agebras of bounded operators only which trivializes domain questions.

Definition III.4.2
The spectrum $\Delta(\mathcal{A})$ of a unital Banach algebra $\mathcal{A}$ is the set of all non-zero $^*-$homomorphisms $\chi :\ \mathcal{A} \to \mathbb{C};\ a \to \chi(a)$, called the characters.

Notice that $\mathbb{C}$ is itself a unital, Abelean C^*-algebra in the usual metric topology of $\mathbb{R}^2$. Notice that $\chi(1) = 1$ since $\chi(a) = \chi(1a) = \chi(1)\chi(a)$ and if we choose $a \in \mathcal{A}$ such that $\chi(a) \neq 0$ the claim follows. Similarly $\chi(a^{-1}) = \chi(a)^{-1}$ if a has an inverse in $\mathcal{A}$, that is an element a^{-1} with $aa^{-1} = a^{-1}a = 1$. Finally $\chi(0) = 0$ since $1 = \chi(1) = \chi(1 + 0) = \chi(1) + \chi(0) = 1 + \chi(0)$.

Definition III.4.3
For a character in a unital Banach algebra $\mathcal{A}$ define $ker(\chi) := \{a \in \mathcal{A};\ \chi(a) = 0\}$ to be its kernel.

Clearly, $\ker(\chi)$ is a two-sided ideal in $\mathcal{A}$ since $\chi(ab) = \chi(ba) = \chi(a)\chi(b) = 0$ for all $a \in \mathcal{A},\ b \in \ker(\chi)$. Since χ is in particular a linear functional on $\mathcal{A}$ considered as a vector space, it follows that $\ker(\chi)$ is a vector subspace of $\mathcal{A}$ of codimension one. After taking its closure in $\mathcal{A}$ it is either still of codimension one or of codimension zero, the latter being impossible since then χ would be identically zero which we excluded in the definition for a character. It follows that there exist elements $a \in \mathcal{A} - \ker(\chi)$ and that $\mathcal{A}$ is the closure of the span of $a, \ker(\chi)$. Thus, if there is an ideal $\mathcal{I}$ of $\mathcal{A}$ properly containing $\ker(\chi)$ then we may take such an $a \in \mathcal{I} - \ker(\chi)$ from which we conclude $\mathcal{I} = \mathcal{A}$. We conclude that the kernel of a character determines a maximal ideal in $\mathcal{A}$.

Definition III.4.4
Let $\mathcal{A}$ be a normed, unital algebra. The spectrum $\sigma(a)$ of $a \in \mathcal{A}$ is defined to be the complement $\mathbb{C} - \rho(a)$ where $\rho(a) := \{z \in \mathbb{C};\ (a - z \cdot 1)^{-1} \in \mathcal{A}\}$ is called the resolvent set of a. For $z \in \rho(a)$ one calls $r_z(a) := (a - z \cdot 1)^{-1}$ the resolvent of $a$at z. The number

$$r(a) := \sup(\{|z|;\ z \in \sigma(a)\} \tag{III.4.2}$$

is called the spectral radius of $a \in \mathcal{A}$.

Notice that the condition $a^{-1} \in \mathcal{A}$ implies that $||a^{-1}||$ exists, that is, the inverse has a norm ("is bounded"). If we are dealing with an algebra of possibly unbounded operators on a Hilbert space then definition III.4.4 must be more precise: if a is a densely defined, closed (the adjoint $a^* \equiv a^\dagger$ is densely defined) linear operator on a Hilbert space $\mathcal{H}$ with dense domain $D(a)$ then $z \in \rho(a)$ iff $a - z \cdot 1$ is a bijection from $D(a)$ onto $\mathcal{H}$ with *bounded* inverse.

We will need later the following technical result.

Lemma III.4.1
For the spectral radius the following identity holds

$$r(a) = \lim_{n \to \infty} ||a^n||^{1/n} \tag{III.4.3}$$

Proof of Lemma III.4.1:
First we show that the series of non-negative numbers $x_n = ||a^n||^{1/n}$ actually converges. For this purpose let $n \geq m \geq 1$ be any natural numbers and split n uniquely as $n = km + r$ for natural numbers k, r with $0 \leq r < m$. By submultiplicativity of the norm we have

$$||a^n||^{1/n} \leq ||a^{km}||^{1/n} \, ||a^r||^{1/n} \leq ||a^m||^{k/n} \, ||a^r||^{1/n} \tag{III.4.4}$$

Fix m and take $n \to \infty$ so that $k = (n-r)/m \to \infty$ while $r \in \{0, .., m-1\}$ stays bounded. Thus the right hand side of (III.4.4) converges to $||a^m||^{1/m}$. It follows that the sequence (x_n), $x_n = ||a^n||^{1/n}$ is bounded and therefore must have an accumulation point each of which must be smaller than x_m for any $m \geq 1$. Let $\lim_n \sup(x_n)$ be the largest accumulation point, then the inequality $\lim_n \sup(x_n) \leq x_m$ holds. Now take the infimum on the right hand side which is also an accumulation point, then we get

$$\lim_n \sup(x_n) \leq \lim_m \inf(x_m) \tag{III.4.5}$$

which means that there is only one accumulation point, so the sequence converges. Denote $x := \lim_{n\to\infty} x_n$.

Now consider the geometrical (von Neumann) series for $z \neq 0$

$$r_z(a) = (a - z \cdot 1)^{-1} = -\frac{1}{z} \sum_{n=0}^{\infty} (\frac{a}{z})^n \tag{III.4.6}$$

which converges if there exists $0 \leq q < 1$ with $||(\frac{a}{z})^n||^{1/n} = ||a^n||^{1/n}/|z| < q$ for all $n > n(q)$. In other words, $z \in \rho(a)$ provided that $|z| > \lim_{n\to\infty} x_n$ or equivalently $z \in \sigma(a)$ provided that

$$|z| \leq x \tag{III.4.7}$$

Taking the supremum in $\sigma(a)$ on the left hand side of (III.4.7) we thus find

$$r(a) \leq x \tag{III.4.8}$$

Suppose now that $r(a) < x$. Then there exists a real number R with $r(a) < R < x$ and since obviously $R \in \rho(a)$ it is clear that the resolvent $r_R(a)$ of a at R converges. Let ϕ be a continuous linear functional on $\mathcal{A}$ then

$$\phi(r_R(a)) = -\frac{1}{R} \sum_{n=0}^{\infty} \phi((\frac{a}{z})^n) \tag{III.4.9}$$

exists which means that $\lim_{n\to\infty} \phi((\frac{a}{z})^n) = 0$. In other words, the function $n \mapsto \phi((\frac{a}{z})^n)$ is bounded for all continuous linear functionals ϕ.

Now the space $\mathcal{A}'$ of continuous linear forms on $\mathcal{A}$ is itself a Banach space with norm $||\phi|| := \sup_{a\in\mathcal{A}} |\phi(a)|$. Consider the family $\mathcal{F} := \{a^n/r^n;\ n \in \mathbb{N}\}$ then we have just shown that for each $b \in \mathcal{F}$ the set $\{|\phi(b)|;\ \phi \in \mathcal{A}'\}$ is bounded. Let us consider each $b \in \mathcal{F}$ as a map $b: \ \mathcal{A}' \to \mathbb{C};\ \phi \to \phi(b)$. We have $||b||' := \sup_{\phi\in\mathcal{A}'} |\phi(b)|/||\phi|| = ||b||$ where the norm in the last equality is the one in $\mathcal{A}$. By the principle of uniform boundedness [129] the set $\{||b||';\ b \in \mathcal{F}\}$ is bounded. Therefore we know that the set of norms $||a^n/r^n||$ is bounded. But

$$||a^n/r^n|| = (\frac{x}{r})^n (\frac{||a^n||^{1/n}}{x})^n \tag{III.4.10}$$

and the first fraction diverges while the second approaches 1 as $n \to \infty$.

Thus in fact $r(a) = \lim_{n\to\infty} ||a^n||^{1/n}$.

□

We will now start establishing the relation between characters and maximal ideals.

Lemma III.4.2
If $\mathcal{I}$ is an ideal in a unital Banach algebra $\mathcal{A}$ then its closure $\overline{\mathcal{I}}$ is still an ideal in $\mathcal{A}$. Every maximal ideal is automatically closed.

Proof of Lemma III.4.2:
Recall that the closure of a subset Y in a topological space is Y together with the limit points of convergent nets in Y. Let now $\mathcal{I}$ be an ideal in $\mathcal{A}$ and let (a^α) be a net in $\mathcal{I}$ converging to $a \in \overline{\mathcal{I}}$. Then for any $b \in \mathcal{A}$ we have $ba^\alpha \in \mathcal{I}$ since $\mathcal{I}$ is an ideal and $\lim_\alpha ba^\alpha = ba$ since $||b(a^\alpha - a)|| \le ||b||\ ||a^\alpha - a|| \to 0$. Thus (ba^α) is a net in $\mathcal{I}$ converging to $ba \in \mathcal{A}$ and since al limit points of converging nets in $\mathcal{I}$ by definition lie in $\overline{\mathcal{I}}$ we actually have have $ba \in \overline{\mathcal{I}}$. Thus, $\overline{\mathcal{I}}$ is an ideal.

Next we notice that every $a \in \mathcal{A}$ such that $||a-1|| < 1$ is invertible (use $a^{-1} = -(1-(a-1))^{-1}$ and the geometric series representation for the latter with convergence radius 1). The set $\{a \in \mathcal{A}; ||a-1|| \ge 1\}$ is a closed subset of $\mathcal{A}$ because if (a^α) is a convergent net in it then the net of real numbers $(||a^\alpha - 1||)$ belongs to the set $\{x \in \mathbb{R};\ x \ge 1\}$ and since it converges to $||a-1||$ it follows that $||a-1|| \ge 1$ since $\{x \in \mathbb{R};\ x \ge 1\}$ is closed (that $b^\alpha \to b$ implies $||b^\alpha|| \to ||b||$ follows from the triangle inequality $||a|| \le ||a-b|| + ||b||, ||b|| \le ||a-b|| + ||a||$). We conclude that every non-trivial (those not containing invertible elements) ideal $\mathcal{I}$ must be contained in the closed set $\{a \in \mathcal{A}; ||a-1|| \ge 1\}$ and so must its closure $\overline{\mathcal{I}}$. Obviously $1 \notin \{a \in \mathcal{A}; ||a-1|| \ge 1\}$, hence, closures of non-trivial ideals are non-trivial.

Finally a maximal ideal must be closed as otherwise its closure would be a non-trivial ideal containing it.
□

Theorem III.4.1 (Gel'fand)
If $\mathcal{A}$ is an Abelean, unital Banach algebra and $\mathcal{I}$ a two-sided, maximal ideal in $\mathcal{A}$ then the quotient algebra $\mathcal{A}/\mathcal{I}$ is isomorphic with $\mathbb{C}$.

Proof of Theorem III.4.1:
By lemma III.4.2 $\mathcal{I}$ is closed in $\mathcal{A}$. We split the proof into three parts.
[i)] *If $\mathcal{I}$ is a maximal ideal in a unital Banach algebra $\mathcal{A}$ then $\mathcal{A}/\mathcal{I}$ is a Banach algebra*
The norm on $\mathcal{A}/\mathcal{I}$ is given by

$$||[a]|| := \inf_{b\in[a]} ||b|| \tag{III.4.11}$$

To see that this indeed defines a norm we check

$$\begin{aligned} ||[za]|| &= ||z[a]|| = \inf_{b\in[a]} ||zb|| = |z|\ ||[a]|| \\ ||[a+a']|| &= ||[a]+[a']|| = \inf_{b\in[a]+[a']} ||b|| = \inf_{b\in[a],b'\in[a']} ||b+b'|| \\ &\le \inf_{b\in[a],b'\in[a']} (||b|| + ||b'||) = ||[a]|| + ||[a']|| \\ ||[a]|| &= \inf_{b\in[a]} ||b|| = 0 \Rightarrow [a] = [0] \end{aligned} \tag{III.4.12}$$

In the second line we exploited that every representative of $[a+a']$ can be written in the form $b+b'$ where b, b' are representatives of $[a], [a']$ and that the joint infimumm is the same as the infimum. The conclusion in the last line means that $[a]$ contains elements of arbitrarily small norm. (Consider a net of elements $(a+b^\alpha)$ in $[a]$ whose norm converges to zero. The net (b^α) is a net in $\mathcal{I}$ and since $\mathcal{I}$ is closed it follows that the limit point $a+b$ lies in $[a]$. Since $||a+b|| = 0$ and $||.||$ is a norm it follows $a+b=0$, thus $0 \in [a]$ and so $[a] = [0]$).

Suppose that $([a_n])$ is a Cauchy sequence in $\mathcal{A}/\mathcal{I}$. We may assume $||[a_{n+1}]-[a_n]|| = ||[a_{n+1}-a_n]|| < 2^{-n}$ (pass to a subsequence if necessary). Since

$$||[a_{n+1}] - [a_n]|| = \inf_{b_{n+1}\in[a_{n+1}],b_n\in[a_n]} ||b_{n+1} - b_n|| < 2^{-n} \tag{III.4.13}$$

we certainly find representatives with $||c_{n+1} - c_n|| < 2^{-n+1}$. Then for $n > m$

$$||c_n - c_m|| = ||\sum_{k=m+1}^{n-1}(c_{k+1} - c_k)|| \le \sum_{k=m+1}^{n-1} 2^{-k+1} = 2^{-m}\sum_{k=0}^{m-n-1} 2^k \le 2^{-m+1} \tag{III.4.14}$$

which displays (c_n) as a Cauchy sequence in $\mathcal{A}$. Since $\mathcal{A}$ is complete this sequence converges to some $a \in \mathcal{A}$. But then

$$||[a_n] - [a]|| = \inf_{b_n\in[a_n],b\in[a]} ||b_n - b|| \le ||c_n - a|| \tag{III.4.15}$$

so $([a_n])$ converges to $[a]$. It follows that $\mathcal{A}/\mathcal{I}$ is complete, that is, a Banach space with unit $[1]$.
$[ii)]$ *For an Abelean, unital algebra $\mathcal{A}$ an ideal $\mathcal{I}$ is maximal in $\mathcal{A}$ iff $\mathcal{A}/\mathcal{I} - [0]$ consists of invertible elements only*
$\Rightarrow$:
Suppose we find $[0] \ne [a] \in \mathcal{A}/\mathcal{I}$ but that $[a]^{-1}$ does not exist. This means that a^{-1} does not exist since $[a]^{-1} = [a^{-1}]$ as follows from $[a][a^{-1}] = [1]$. Consider now the ideal $\mathcal{A}\cdot a = \{ba;\ b \in \mathcal{A}\}$ (this is a two-sided ideal because $\mathcal{A}$ is Abelean). Since $\mathcal{I} \subset \mathcal{A}$ we certainly have $\mathcal{I}\cdot a \subset \mathcal{A}\cdot a$ and since $\mathcal{I}\cdot a = \mathcal{I}$ because $\mathcal{I}$ is in particular a right ideal we have $\mathcal{I} \subset \mathcal{A}\cdot a$. Now $a \in \mathcal{A}\cdot a$ since $1 \in \mathcal{A}$ and $a \notin \mathcal{I}$ because otherwise $[a] = [0]$ which we excluded. It follows that $\mathcal{I}$ is a proper subideal of $\mathcal{A}\cdot a$. Finally, since $a^{-1} \notin \mathcal{A}$, $\mathcal{A}\cdot a$ cannot be all of $\mathcal{A}$, for instance $1 \notin \mathcal{A}\cdot a$ (an ideal that contains 1 or any invertible element is anyway the whole algebra). It follows that $\mathcal{I}$ is not maximal.
$\Leftarrow$:
Suppose $\mathcal{I}$ is not a maximal ideal. Then we find a proper subideal $\mathcal{J}$ of $\mathcal{A}$ of which $\mathcal{I}$ is a proper subideal. Since every non-zero element of $\mathcal{A}/\mathcal{I}$ is invertible so is every element $[a]$ of $\mathcal{J}/\mathcal{I}$. But then $\mathcal{J}$ contains the invertible element $a \in \mathcal{A}$ and thus $\mathcal{J}$ coincides with $\mathcal{A}$ which is a contradiction.
$[iii)]$ *A unital Banach algebra $\mathcal{B}$ in which every non-zero element is invertible is isomorphic with $\mathbb{C}$*
Consider any $b \in \mathcal{B}$ then we claim that $\sigma(b) \ne \emptyset$. Suppose that were not the case then $\rho(b) = \mathbb{C}$. Let ϕ be a continuous linear functional on $\mathcal{A}$ considered as a vector space with metric. Using linearity of ϕ and the expansion of $r_z(b)$ into an absolutely geometric series we see that $z \mapsto \phi(r_z(b))$ is an entire analytic function. Since ϕ is linear and continuous, it is bounded with bound $||\phi||$. Thus $|\phi(r_z(b))| \le ||\phi||\ ||r_z(b)||$. Since $\lim_{z\to\infty} ||r_z(b)|| = 0$ (use the geometric series) and $||r_z(a)||$ is everywhere defined in $\mathbb{C}$ we conclude that $z \mapsto \phi(r_z(b))$ is an entire bounded function which therefore, by Liouville's theorem, is a constant $c_a = \phi(r_z(b)) = \lim_{z\to\infty}\phi(r_z(b)) = 0$. Since ϕ was arbitrary it follows that $r_z(a) = 0$ implying that $b - z\cdot 1$ does not exist which cannot be the case.

Thus we find $z_b \in \sigma(b)$, that is, $b - z_b\cdot 1$ is not invertible. By assumption, only zero elements are not invertible, hence $b = z_b \cdot 1$ for some $z_b \in \mathbb{C}$ for any $b \in \mathcal{B}$. The map $b \mapsto z_b$ is then the searched for isomorphism $\mathcal{B} \to \mathbb{C}$. Notice that $b = 0$ iff $z_b = 0$.

Let then $\mathcal{I}$ be a maximal ideal in a unital, Abelean Banach algebra $\mathcal{A}$. Then by i) $\mathcal{B} := \mathcal{A}/\mathcal{I}$ is a unital Banach algebra and by ii) each of its non-zero elements is invertible. Thus by iii) it is isomorphic with $\mathbb{C}$.
□

Corollary III.4.1
In an Abelean, unital Banach algebra $\mathcal{A}$ there is a one-to-one correspondence between its spectrum $\Delta(\mathcal{A})$ and the set $I(\mathcal{A})$ of maximal ideals in $\mathcal{A}$ via

$$\Delta(\mathcal{A}) \to I(\mathcal{A});\ \chi \mapsto ker(\chi) \tag{III.4.16}$$

Proof of Corollary III.4.1:
That each character gives rise to a maximal ideal in $\mathcal{A}$ through its kernel was already shown after definition III.4.3. Conversely, let $\mathcal{I}$ be a maximal ideal in a commutative unital Banach algebra then we can apply theorem III.4.1 and obtain a Banach algebra isomorphism $\chi : \mathcal{A}/\mathcal{I} \to \mathbb{C};\ [a] \to \chi([a])$. We can extend this to a homomorphism $\chi :\ \mathcal{A} \to \mathbb{C}$ by $\chi(a) := \chi([a])$. By construction $\chi(a) = 0$ iff $[a] = [0]$, that is, iff $a \in \mathcal{I}$. In other words, the maximal ideal $\mathcal{I}$ is the kernel of the character χ.
□
The subsequent lemma explains the word "spectrum".

Lemma III.4.3
Let $\mathcal{A}$ be a unital, commutative Banach algebra and $a \in \mathcal{A}$. Then $z \in \sigma(a)$ iff there exists $\chi \in \Delta(\mathcal{A})$ such that $\chi(a) = z$.

Proof of Lemma III.4.3:
The requirement $\chi(a) = z$ is equivalent with $\chi(a - z \cdot 1) = 0$ so that $a - z \cdot 1 \in \ker(\chi)$. Since $\ker(\chi)$ is a maximal ideal in $\mathcal{A}$ it cannot contain invertible elements, thus $(a - z \cdot 1)^{-1}$ does not exist, hence $z \in \sigma(a)$.
□
We now equip the spectrum with a topology. We begin by showing that the characters are in particular continuous linear functionals on the topological vector space $\mathcal{A}$.

Definition III.4.5
For a character χ in an Abelean, unital Banach algebra we define its norm by

$$||\chi|| := \sup_{a\in\mathcal{A}} |\chi(a)| \tag{III.4.17}$$

Lemma III.4.4
The characters of an Abelean, unital Banach algebra form a subset of the unit sphere in $\mathcal{A}'$, the continuous linear functionals on $\mathcal{A}$ considered as a topological vector space.

Proof of Theorem III.4.4:
By lemma III.4.3 we showed that $\sigma(a) = \{\chi(a);\ \chi \in \Delta(\mathcal{A})\}$. It follows that

$$||\chi|| = \sup_{a\in\mathcal{A}} \frac{|\chi(a)|}{||a||} \leq \sup_{a\in\mathcal{A}} \frac{\sup\{|\chi'(a)|;\ \chi' \in \Delta(\mathcal{A})\}}{||a||} = \sup_{a\in\mathcal{A}} \frac{\rho(a)}{||a||} \leq 1 \tag{III.4.18}$$

since by lemma III.4.1 we have $r(a) = \lim_{n\to\infty} ||a^n||^{1/n} \leq ||a||$. On the other hand $\chi(1) = 1$, hence $||\chi|| = 1$ for every character χ. This shows that every character is a bounded linear functional on $\mathcal{A}$, that is, $\Delta(\mathcal{A}) \subset \mathcal{A}'$.
□
Since we just showed that the characters are in particular bounded linear functionals it is natural to equip the spectrum with the weak * topology of pointwise convergence induced from $\mathcal{A}'$.

Definition III.4.6
i)
*The weak * topology on the topological dual X' of a topological vector space X (the set of continuous (bounded) linear functionals) is defined by pointwise convergence, that is, a net (ϕ^α) in X' converges to ϕ iff for any $x \in X$ the net of complex numbers $(\phi^\alpha(x))$ converges to $\phi(x)$). Equivalently, it is the weakest topology such that all the functions $x:\ X' \to \mathbb{C}; \phi \to \phi(x)$ are continuous.*
ii)
*The Gel'fand topology on the spectrum of a unital, Abelean Banach algebra is the weak * topology induced from $\mathcal{A}'$ on its subset $\Delta(\mathcal{A})$.*

We now show that in the Gel'fand topology the spectrum becomes a compact Hausdorff space. We need a preparational lemma.

Lemma III.4.5
*Let X be a Banach space and X' its topological dual. Then the unit ball in X' is closed and compact in the weak * topology.*

Proof of Lemma III.4.5:
The unit ball B in X' is defined as the subset of elements ϕ with norm smaller than or equal to unity, that is, $||\phi|| := \sup_{x\in X} |\phi(x)|/||x|| \le 1$. By corollary III.3.1 we must show that every universal net in B converges. Let ϕ^α be a universal net in B and consider for any given $x \in X$ the net of complex numbers $(\phi^\alpha(x))$ which are bounded by $||x||$. Our $x \in X$ defines a linear form $X' \to \mathbb{C};\ \phi \to \phi(x)$ whence by lemma III.3.2ii) the net $(\phi^\alpha(x))$ is universal. It is contained in the set $\{z \in \mathbb{C};\ |z| \le ||x||\}$ which is compact in $\mathbb{C}$ and therefore it converges. Define ϕ pointwise by the limit, that is, $\phi(x) := \lim_\alpha \phi^\alpha(x)$. Then

$$||\phi|| = \sup_{x\in X} \lim_\alpha |\phi^\alpha(x)|/||x|| \le \lim_\alpha ||\phi^\alpha|| \le 1 \tag{III.4.19}$$

Thus ϕ^α converges pointwise to $\phi \in B$. In particular we have shown that B is closed.
□

Theorem III.4.2
In the Gel'fand topology, the spectrum $\Delta(\mathcal{A})$ of a unital, Abelean Banach algebra is compact.

Proof of Theorem III.4.2:
Since we have shown 1) in lemma III.4.4 that $\Delta(\mathcal{A})$ is a subset of the unit ball B in $\mathcal{A}'$, 2) in lemma III.4.5 that B is compact in the weak * topology and 3) in lemma III.3.3 that closed subspaces of compact spaces are are compact in the subspace topology it will be sufficient to show that $\Delta(\mathcal{A})$ is a closed in B as the Gel'fand topology is the subspace topology induced from B.

Let then (χ^α) be a net in $\delta(\mathcal{A})$ converging to $\chi \in B$. We have, e.g., $\chi(ab) = \lim_\alpha \chi^\alpha(ab) = \lim_\alpha \chi^\alpha(a)\chi^\alpha(b) = \chi(a)\chi(b)$ and similar for pointwise addition, scalar multiplication and involution in $\mathcal{A}$. It follows that χ is a character, that is, $\chi \in \Delta(\mathcal{A})$.
□

Definition III.4.7
The Gel'fand transform is defined by

$$\bigvee:\ \mathcal{A} \to \Delta(\mathcal{A})';\ a \mapsto \check{a}\ \textit{where}\ \check{a}(\chi) := \chi(a) \tag{III.4.20}$$

Here $\Delta(\mathcal{A})'$ denotes the continuous linear functionals on $\Delta(\mathcal{A})$ considered as a topological vector space.

It is clear that every $\check{a},\ a \in \mathcal{A}$ is a continuous linear functional on the spectrum since for any net (χ^α) in $\Delta(\mathcal{A})$ which converges to χ we have $\lim_\alpha \check{a}(\chi^\alpha) = \lim_\alpha \chi^\alpha(a) = \chi(a) = \check{a}(\chi)$ because convergence of (χ^α) means pointwise convergence on $\mathcal{A}$.

Theorem III.4.3
The Gel'fand transform extends to a homomorphism

$$\bigvee :\ \mathcal{A} \to C(\Delta(\mathcal{A}));\ a \to \check{a} \tag{III.4.21}$$

with the following additional properties:
1) range$(\check{a}) = \sigma(a)$.
2) $||\check{a}|| := \sup_{\chi\in\Delta(\mathcal{A})} |\check{a}(\chi)| = r(a)$.
3) The image $\bigvee(\mathcal{A})$ separates the points of $\Delta(\mathcal{A})$.

Proof of Theorem III.4.3:
0)
Morphism and Continuity:
We have for example

$$(ab)^{\bigvee}(\chi) = \chi(ab) = \chi(a)\chi(b) = \check{a}(\chi)\check{b}(\chi) \tag{III.4.22}$$

for any $\chi \in \Delta(\mathcal{A})$ and similar for $(a+b)^{\bigvee}$. Thus multiplication and addition of functions are defined pointwise. That the functions $\check{a}$ are continuous follows as after definition III.4.7 from the fact that the weak * topology on $\Delta(\mathcal{A})$ *is defined* by asking that all the Gel'fand transforms $\check{a}$ be continuous and therefore is tautologous.
1)
We have

$$\text{range}(\check{a}) = \{\check{a}(\chi);\ \chi \in \Delta(\mathcal{A})\} = \{\chi(a);\ \chi \in \Delta(\mathcal{A})\} = \sigma(a) \tag{III.4.23}$$

as follows from lemma III.4.3.
2)
We have

$$||\check{a}|| = \sup_{\chi\in\Delta(\mathcal{A})} |\check{a}(\chi)| = \sup_{\chi\in\Delta(\mathcal{A})} |\chi(a)| = \sup(\{|\chi(a)|;\ \chi \in \Delta(\mathcal{A})\}) = r(a) \tag{III.4.24}$$

by definition of the spectral radius. Notice that the sup-norm is a natural norm on a space of continuous functions on a compact space.
3)
Recall that a collection of functions $\mathcal{C}$ on a topological space X is said to separate its points iff for any $x_1 \neq x_2$ we find $f \in \mathcal{C}$ such that $f(x_1) \neq f(x_2)$. Consider then any $\chi_1, \chi_2 \in \Delta(\mathcal{A})$ with $\chi_1 \neq \chi_2$. By definition of $\Delta(\mathcal{A})$ there exists then $a \in \mathcal{A}$ such that $\chi_1(a) = \check{a}(\chi_1) \neq \chi_2(a) = \check{a}(\chi_2)$.
$\Box$
To see that then $\Delta(\mathcal{A})$ is a Hausdorff space recall the following lemma.

Lemma III.4.6 *Let X be a topological space and $\mathcal{C} \subset C(X)$ a collection of continuous functions on X which separate the points of X. Then the topology on X is Hausdorff.*

Proof of Lemma III.4.6:
Let $x_1, x_2 \in X$ with $x_1 \neq x_2$ be any two distinct points. Since $\mathcal{C}$ separates the points we find $f \in \mathcal{C}$ with $f(x_1) \neq f(x_2)$. Let $d := |f(x_2) - f(x_1)|$. Since f is continuous at x_I, for any $\epsilon > 0$ we find a neighbourhood $U_I(\epsilon)$ of x_I, $I = 1, 2$ such that $|f(x) - f(x_I)| < \epsilon$ for any $x \in U_I(\epsilon)$. Now $d = |f(x_2) - f(x_1)| \leq |f(x) - f(x_1)| + |f(x_2) - f(x)|$ for any $x \in X$. Thus $d - \epsilon < |f(x_2) - f(x)|$ for any $x \in U_1(\epsilon)$ and $d - \epsilon < |f(x_1) - f(x)|$ for any $x \in U_2(\epsilon)$. Choose $\epsilon < d/2$. Then $U_1(\epsilon) \cap U_2(\epsilon) = \emptyset$.
□

Corollary III.4.2
The Gel'fand topology on the spectrum of a unital, Abelean Banach algebra is Hausdorff.

Proof of Corollary III.4.2:
The proof follows trivially from the fact that by theorem III.4.3 $\mathcal{C} := \{\check{a};\ a \in \mathcal{A}\}$ is a system of continuous functions separating the points of $\Delta(\mathcal{A})$ together with lemma III.4.6.
□
So far everything worked for an Abelean, unital Banach algebra $\mathcal{A}$. We now invoke the further restriction that $\mathcal{A}$ be an Abelean, unital C^* algebra which makes the Gel'fand transform especially nice.

Theorem III.4.4
Let $\mathcal{A}$ be a unital, commutative C^*-algebra (not only a Banach algebra). Then the Gel'fand transform is an isometric isomorphism between $\mathcal{A}$ and the space of continuous functions on its spectrum.

Proof of Theorem III.4.4:
First of all, using the fact that in a commutative * algebra every element is normal (meaning that $[a,^*] = 0$) we have, making frequent use of the C^* property (III.4.1)

$$\begin{aligned} ||a^{2^n}||^2 &= ||a^{2^n}(a^{2^n})^*|| = ||(aa^*)^{2^n}|| \\ &= ||(aa^*)^{2^{n-1}}((aa^*)^{2^{n-1}})^*|| = ||(aa^*)^{2^{n-1}}||^2 \\ &= ||aa^*||^{2^n} = ||a||^{2^{n+1}} \end{aligned} \tag{III.4.25}$$

where in the third equality we exploited that aa^* is self-adjoint an in the fifth equality we iterated the equality between the expressions at the end of the first and second line. We conclude that for any natural number n

$$||a|| = ||a^{2^n}||^{1/2^n} \tag{III.4.26}$$

In lemma III.4.1 we proved the formula $r(a) = \lim_{n\to\infty} ||a^n||^{1/n}$ meaning that every subsequence of the sequence $(||a^n||^{1/n})$ has the same limit $r(a)$ including the one displayed in (III.4.26). Thus we have shown that for Abelian C^*-algebras indeed

$$r(a) = ||a|| \tag{III.4.27}$$

and not only $r(a) \leq ||a||$. By theorem III.4.32) we have therefore

$$||\check{a}|| = ||a|| \tag{III.4.28}$$

that is, isometry.

Consider now the system of complex valued functions on the spectrum given by $\mathcal{C} := \{\check{a};\ a \in \mathcal{A}\}$. We claim that it has the following properties:

i) $\mathcal{C} \subset C(\Delta(\mathcal{A}))$
ii) $\mathcal{C}$ separates the points of $\Delta(\mathcal{A})$
iii) $\mathcal{C}$ is a closed (in the sup-norm topology) * subalgebra of $C(\Delta(\mathcal{A}))$
iv) The constant functions belong to $\mathcal{C}$.
Property i), ii) are the assertions 0) and 3) of theorem III.4.3 while iv) follows from the fact that $\mathcal{A}$ is unital, i.e. $\check{1}(\chi) = \chi(1) = 1$ so $\check{1} = 1$. To show that iii) $\mathcal{C}$ is a closed * algebra in $C(\Delta(\mathcal{A}))$ suppose that $(\check{a}^\alpha)$ is a net in $\mathcal{C}$ converging to some $f \in C(\Delta(\mathcal{A}))$. Thus, $(\check{a}^\alpha)$ is in particular a Cauchy sequence, meaning that $||\check{a}^\alpha - \check{a}^\beta|| = ||a^\alpha - a^\beta||$ becomes arbitrarily small as $\alpha\beta$ grow, where we have used isometry. It follows that (a^α) is a Cauchy seqence and therefore converges to some $a \in \mathcal{A}$ since $\mathcal{A}$ is in particular a Banach algebra and therefore complete. Therefore $f = \check{a} \in \mathcal{C}$, whence $\mathcal{C}$ is closed. Clearly $\mathcal{C}$ is also a * subalgebra because $\mathcal{A}$ is an algebra and $\bigvee$ a homomorphism.

Now reacll from theorem III.4.2 and corollary III.4.2 that $\Delta(\mathcal{A})$ is a compact Hausdorff space. Then properties i),ii),iii) of $\mathcal{C}$ enable us to apply the Stone-Weierstrass theorem (e.g. [129]) which tells us that either $\mathcal{C} = C(\Delta(\mathcal{A}))$ or that there exists $\chi_0 \in \Delta(\mathcal{A})$ such that $\check{a}(\chi_0) = 0$ for all $\check{a} \in \mathcal{C}$. By properety iv) the latter possibility is excluded whence $\mathcal{C} = \bigvee(\mathcal{A})$ is *all of* $C(\Delta(\mathcal{A}))$. In other words, the Gel'fand transform is a surjection. Finally it is an injection since $\check{a} = \check{a}'$ implies $||\check{a} - \check{a}'|| = ||a - a'|| = 0$ by isometry, hence $a = a'$.
□

Corollary III.4.3
Every compact Hausdorff space X arises as the spectrum of an Abelean, unital C^*- algebra $\mathcal{A}$, specifically $\mathcal{A} = C(X)$, $\Delta(\mathcal{A}) = X$.

Proof of Corollary III.4.3:
Let X be a compact Hausdorff space and define $\mathcal{A} := C(X)$ equipped with the sup-norm. Then $X \subset \Delta(C(X))$ by the defintion $x(f) := f(x) =: \check{f}(x)$ for any $f \in \mathcal{A}$ so the Gel'fand transform is the identity map on X. Thus, if $\Delta(C(X)) - X \neq \emptyset$ then $\check{f}$ extends f continuously to $\Delta(C(X))$.

Next let (x^α) be a net in X which converges in $\Delta(C(X))$ then $\check{f}(x^\alpha)$ converges in $\mathbb{C}$ for any $\check{f} \in C(\Delta(C(X)))$, i.e., $f(x^\alpha)$ converges in $\mathbb{C}$ for any $f \in C(X)$. It follows that (x^α) converges in X, that is, X is closed in $\Delta(C(X))$.

Suppose now that $\Delta(C(X)) - X \neq \emptyset$. Thus we find $\chi_0 \in \Delta(C(X)) - X$. Now in a Hausdorff space the one point sets are closed [140]. Therefore the sets $X, \{\chi_0\}$ are disjoint closed sets in the compact Hausdorff space $\Delta(C(X))$. Since compact Hausdorff spaces are normal spaces [129] (i.e. one point sets are closed and any two disjoint closed sets are contained in open disjoint sets) we may apply Urysohns's lemma [129] to conclude that there is a continuous function $F : \Delta(C(X)) \to \mathbb{R}$ with range in $[0,1]$ such that $F_{|X} = 0$ and $F|\{\chi_0\} = F(\chi_0) = 1$.

Consider then any $f \in C(X)$. Since $C(\Delta(C(X)))$ are *all* continuous functions on $\Delta(C(X))$, there exist different continuous extensions of f to $\Delta(C(X))$, for instance the functions $\check{f}, \check{f} + F$ where F is of the form just constructed. However, this contradicts the fact that $\bigvee$ is an isomorphism since it would not be surjective.
□
Corollary III.4.3 tells us that a compact Hausdorff space can be reconstructed from its Abelean, unital C^*-algebra of continuous functions by constructing its spectrum. This is the starting point for generalizations to non-commutative topological spaces [39].

III.5 Tools from Measure Theory

For an introduction to general measure theory see e.g. the beautiful textbook [145] . For more advanced topics concerning the extension theory of measures from self-consistent families of projections to $\sigma-$additive ones, see e.g. [58].

Recall the notion of a topology and of continuous functions from section III.3.

Definition III.5.1
i)
Let X be a set. Then a collection of subsets $\mathcal{U}$ of X is called a $\sigma-$algebra provided that
1) $X \in \mathcal{U}$,
2) $U \in \mathcal{U}$ implies $X - U \in \mathcal{U}$ and
3) $\mathcal{U}$ is closed under countabe unions, that is, if $U_n \in \mathcal{U},\ n = 1, 2, ..$ then also $\cup_{n=1}^{\infty} U_n \in \mathcal{U}$.
The sets $U \in \mathcal{U}$ are called measurable and a space X equipped with a $\sigma-$algebra a measurable space.
ii)
Let X be a measurable space and let Y be a topological space. A function $f : X \to Y$ is said to be measurable provided that the preimage $f^{-1}(V) \subset X$ of any open set $V \subset Y$ is a measurable subset in X.
iii)
Let X be a topological space. The smallest $\sigma-$algebra on X that contains all open (and due to 2) therefore all closed) sets of X is called the Borel $\sigma-$algebra of X. The elements of the Borel $\sigma-$algebra are called Borel sets.

Given a collection $\mathcal{U}$ of subsets of X which is not yet a topology ($\sigma-$algebra) the weakest topology (smallest $\sigma-$algebra) containing $\mathcal{U}$ is obtained by adding to the collection the sets $X, \emptyset$ as well as arbitrary unions plus finite intersections (countable unions and intersections). Notice the similarity between a collection of sets $\mathcal{U}$ that qualify for a $\sigma-$algebra and a topology: In both cases the sets $X, \emptyset$ belong to $\mathcal{U}$ but while open sets are closed under arbitrary unions and finite intersections, measurable sets are closed under countable unions and intersections. Note also that if X, Y are topological spaces and $f : X \to Y$ is continuous then f is automatically measurable if X is equipped with the Borel $\sigma-$algebra.

Definition III.5.2
A complex measure μ on a measurable space $(X, \mathcal{U})$ is a function $\mu : \mathcal{U} \to \mathbb{C} - \infty;\ U \mapsto \mu(U)$ which is countably (or $\sigma-$)additive, that is,

$$\mu(\bigcup_{n=1}^{\infty} U_n) = \sum_{n=1}^{\infty} \mu(U_n) \tag{III.5.1}$$

for any mutually disjoint measurable sets U_n. A positive measure is also a $\sigma-$additive map $\mu : \mathcal{U} \to \mathbb{R}^+ \cup 0, \infty$ which however is postive semidefinite and may take the value ∞ with the convention $0 \cdot \infty = 0$ (which makes $[0, \infty]$ a set in which commutative, distributive and associative law hold). To avoid trivialities we assume that $\mu(U) < \infty$ for at least one measurable set U. A measure is called a probability measure if $\mu(X) = 1$. The triple $(X, \mathcal{U}, \mu)$ is called a measure space.

In what follows we will always assume that μ is a positive measure.
A very powerful tool in measure theory are characteristic functions of subsets of X.

Definition III.5.3
A function $s: X \to \mathbb{C}$ in a measurable space $(X,\mathcal{U})$ is called simple provided its range consists of finitely many points only. If $z_k \in \mathbb{C}$, $k = 1,..,N$ are these values and $S_k = s^{-1}(\{z_k\})$ then $s = \sum_{k=1}^N z_k \chi_{S_k}$ where χ_S with ($\chi_S(x) = 1$ if $x \in S$ and $\chi_S(x) = 0$ otherwise) is called the characteristic function of the subset $S \subset X$. Obviously, a simple function is measurable if and only if the S_k are measurable.

The justification for this definition lies in the following lemma.

Lemma III.5.1
Let $f: X \to [0,\infty]$ be measurable. Then there exists a sequence of measurable simple functions s_n such that
a) $0 \le s_1 \le s_2 \le ... \le f$
b) $\lim_{n=1} s_n(x) = f(x)$ *pointwise in* $x \in X$.

The proof can be found in [145], theorem 1.17.

Definition III.5.4 *i)*
For a simple measurable function $s = \sum_{k=1}^N z_k \chi_{S_k}$ with $z_k > 0$ on a measure space $(X,\mathcal{U},\mu)$ with positive measure μ we define

$$\mu(s) := \int_X d\mu(x) s(x) := \sum_{k=1}^N z_k \mu(S_k) \tag{III.5.2}$$

For a general measurable function $f: X \to [0,\infty]$ we define

$$\mu(f) := \sup_{0 \le s \le f} \mu(s) \tag{III.5.3}$$

where the supremum is taken over the simple, positive measurable functions that are nowhere larger than f. The number $\mu(f)$ is called the Lebesgue integral of f. For a general complex valued, measurable function f one can show that we have a unique split as $f = u + iv$, $u = u_+ - u_-$, $v = v_+ - v_-$ with non-negative measurable functions $u_\pm, v_\pm$ and the integral is defined as $\mu(f) = \mu(u_+) - \mu(u_-) + i[\mu(u_+) - \mu(u_-)]$. Also $|f|$ can be shown to be measurable.
ii)
A measure μ is called positive definite if for every non-negative measurable function f the condition $\mu(f) = 0$ implies $f = 0$ almost everywhere (a.e., i.e. up to measure zero sets).

Of fundamental importance are conditions under which one is allowed to echange integration and taking limits.

Theorem III.5.1
Let $(X,\mathcal{U},\mu)$ be a measure space with positive measure μ and let (f_n) be a sequence of measurable functions that converges pointwise on X to the function f.
i) Lebesgue Monotone Convergence Theorem
Suppose that $0 \le f_n(x) \le f_{n+1}(x)$ for all $x \in X$. Then f is measurable and $\lim_{n\to\infty} \mu(f_n) = \mu(f)$.
ii) Lebesgue Dominated Convergence Theorem
A function F is said to be in $L_1(X, d\mu)$ if it is measurable and $\mu(|F|) < \infty$. Suppose now that there exists $F \in L_1(X, s\mu)$ such that $|f_n(x)| \le |F(x)|$ for all $x \in X$. Then $f \in L_1(X, d\mu)$ and $\lim_{n\to\infty} \mu(|f - f_n|) = 0$.

It is easy to see that $\lim_{n\to\infty}\mu(|f-f_n|)=0$ implies $\lim_{n\to\infty}\mu(f_n)=\mu(f)$.
Another convenient observation is the following.

Theorem III.5.2
Let $(X,\mathcal{U},\mu)$ be a measure space. Let $\mathcal{U}'$ be the collection of all $S\subset X$ such that there exist $U,V\in\mathcal{U}$ with $U\subset S\subset V$ and $\mu(V-U)=0$ (in particular $\mathcal{U}\subset\mathcal{U}'$). Define $\mu'(S)=\mu(U)$ in that case. Then $(X,\mathcal{U}',\mu')$ is a measure space again, called the completion of $(X,\mathcal{U},\mu)$.

The theorem says that any measure can be completed. It means that if we have a set which is not measurable but which can be sandwiched between measurable sets whose difference has zero measure, then we can add the set to the measurable sets and its measure is given by that of the sandwiching sets.

Definition III.5.5
i)
A set $Y\subset X$ in a measure space $(X,\mathcal{U},\mu)$ is called thick or a support for μ provided that for any measurable set $U\in\mathcal{U}$ the condition $U\cap Y=\emptyset$ implies $\mu(U)=0$. A support for μ will be denoted by $supp(\mu)$.
ii)
For two measures μ_1,μ_2 on the same measurable space we say that μ_1 is regular (or absolutely continuous) with respect to μ_2 iff $\mu_2(U)=0$ for $U\in\mathcal{U}$ implies $\mu_1(U)=0$. They are called mutually singular iff $supp(\mu_1)\cap supp(\mu_2)=\emptyset$.

If Y is a measurable support then $X-Y$ is measurable and since $Y\cap(X-Y)=\emptyset$ we have $\mu(X-Y)=0$ explaining the word support. If Y is a support not measurable with respect to μ one can define $\mathcal{U}'=[\mathcal{U}\cap Y]\cup Y$, $\mu'(U\cap Y)=\mu(U)$ and gets a measure space $(Y,\mathcal{U}',\mu')$ for which Y is measurable, called the trace. A given support does not mean that there are not smaller sets which are still thick. If μ_2 is a positive $\sigma-$finite (see below) measure and μ_1 is a complex measure, then one can show (the Radon-Nikodym theorem) that there is a unique (so-called Lebesgue) decomposition $\mu_1=\mu_1^a+\mu_1^s$ such that μ_1^a,μ_2^s are repectively absolutely continuous and singular with respect to μ_2 and that there exists $f\in L_1(X,d\mu_2)$, called the Radon-Nikodym derivative, such that $d\mu_1^a=f\;d\mu_2$.

The following two definitions prepares to state the Riesz representation (or Riesz – Markov) theorem which will be of fundamental importance for our applications.

Definition III.5.6
i) A topological space is said to be locally compact if every point $x\in X$ has an open neighbourhood whose closure is compact.
ii)
A subset $S\subset X$ of a topological space X is said to be $\sigma-$compact if it is a countable union of compact sets.
iii)
A subset $S\subset X$ in a measure space $(X,\mathcal{U},\mu)$ with positive measure μ is said to be $\sigma-$finite if S is the countable union of measurable sets U_n with $\mu(U_n)<\infty$ for all $n\in\mathbb{N}$.

Definition III.5.7
Let X be a locally compact Hausdorff space and let $\mathcal{U}$ be its naturally defined Borel $\sigma-$algebra.
i)
A measure μ defined on the Borel $\sigma-$algebra is called a Borel measure.

ii)
A Borel set S is said to be outer regular with respect to a positive Borel measure μ provided that

$$\mu(S) = \inf\{\mu(O);\ S \subset O;\ O \in \mathcal{U}\ open\} \tag{III.5.4}$$

iii)
A Borel set S is said to be inner regular with respect to a positive Borel measure μ provided that

$$\mu(S) = \sup\{\mu(K);\ S \supset K;\ K \in \mathcal{U}\ compact\} \tag{III.5.5}$$

iv) If μ is a positive Borel measure and every Borel set is both inner and outer regular then μ is called regular.

Definition III.5.8
i)
Let X be a topological space. The support supp(f) of a function $f:\ X \to \mathbb{C}$ is the closure of the set $\{x \in;\ f(x) \neq 0\}$. The vector space of continuous functions of compact support is denoted by $C_0(X)$.
ii)
A linear functional $\Lambda:\ \mathcal{F} \to \mathbb{C}$ on the vector space of functions $\mathcal{F}$ over a set X is called positive if $\Lambda(f) \in [0,\infty)$ for any $f \in \mathcal{F}$ such that $f(x) \in [0,\infty)$ for all $x \in X$.

Theorem III.5.3 (Riesz Representation Theorem)
i)
Let X be a locally compact Hausdorff space and let $\Lambda:\ C_0(X) \to \mathbb{C}$ be a positive linear functional one the space of continuous, complex-valued functions of compact support in X. Then there exists a $\sigma-$algebra $\mathcal{U}$ on X which contains the Borel $\sigma-$algebra and a unique positive measure μ on $\mathcal{U}$ such that Λ is represented by μ, that is,

$$\Lambda(f) = \int_X d\mu(x) f(x) \quad \forall\, f \in C_0(X) \tag{III.5.6}$$

Moreover, μ has the following properties:
1) $\mu(K) < \infty$ if $K \subset X$ is compact.
2) For every $S \in \mathcal{U}$ property (III.5.4) holds.
3) For every open $S \in \mathcal{U}$ with $\mu(S) < \infty$ property (III.5.5) holds.
4) If $S' \subset S \in \mathcal{U}$ and $\mu(S) = 0$ then $S' \in \mathcal{U}$.
ii)
If, in addition to i), X is $\sigma-$compact then μ has the following additional properties:
5) μ is regular
6) For any $S \in \mathcal{U}$ and any $\epsilon > 0$ there exists a closed set C and an open set O such that $C \subset S \subset O$ and $\mu(O - C) < \epsilon$.
7) For any $S \in \mathcal{U}$ there exist sets C' and O' which are respectively countable unions and intersections of closed and open sets respectively such that $C' \subset S \subset O'$ and $\mu(O' - C') = 0$.

A very instructive proof of this theorem can be found in [145]. It is also worth pointing out the following theorem (see e.g. [145]) which underlines the prominent role that continuous functions play for Borel measures.

Theorem III.5.4 (Lusin's Theorem)
Let X be a locally compact Hausdorff space X with $\sigma-$algebra $\mathcal{U}$ and measure μ satisfying the properties 1), 2), 3) and 4) of theorem III.5.3. Let f be a bounded measurable function with support in a measurable set of finite measure. Then there exists a sequence (f_n) of continous functions of compact support, each of which is bounded by the same bound, such that $f(x) = \lim_{n\to\infty} f_n(x)$ almost everywhere with respect to μ (i.e. they coincide pointwise up to sets of measure zero).

Let us also define the notion of faithfulness of measures:

Definition III.5.9
Let X be a locally compact Hausdorff space and let $\mathcal{U}, \mu$ have the properties of theorem III.5.3. Then μ is called faithful if and only if the positive linear functional (III.5.6) determined by μ is positive definite, that is, if $f \in C_0(X)$ takes only values in $[0,\infty)$ and $\Lambda(f) = 0$ then $f = 0$.

Notice that positive definiteness of a measure μ only allows us to conclude that $f = 0$ $\mu-$a.e. from $\mu(f) = 0$ for positive measurable f. Faithfulness of the special kind of measures that come from positive definite linear functionals alow us to conclude $f = 0$ everywhere if f is continuous and of compact support. This means that every open set must have positive measure for if a continuous function is positive at a point, it will be bounded away from zero in a whole open neighbourhood of that point.

The application that we have in mind is that X is not only locally compact but actually compact so that the set $C_0(X)$ coincides with $C(X)$. Hence, $C(X)$ contains the constant functions and we may w.l.g. assume that $\Lambda(1) = 1$ which is just a convenient choice of normalization. (If X is compact, so is every closed subset, hence X is locally compact). It is then trivially $\sigma-$compact being its own cover by compact sets. Therefore the stronger version ii) of theorem III.5.3 applies and we see that by property 5) the measure μ is regular. Furthermore, property 7) tells us that every measurable set can be sandwiched between sets $C' \subset O'$ that belong to the Borel $\sigma-$subalgebra such that $C' - O'$ is of measure zero. In other words, every measurable set is a Borel set up to a set of measure zero: Since $O' = S \cup (O' - S)$ we have from $\sigma-$additivity $\mu(S) = \mu(O')$ since $0 = \mu(O' - C') \geq \mu(O' - S)$ due to $O' - S \subset O' - C'$. Thus effectively the measure μ in (III.5.6) is a Borel measure and in that sense we have the following corollary.

Corollary III.5.1
Let X be a compact Hausdorff space and let $\Lambda : C(X) \to \mathbb{C}$ be a positive linear functional on the space of continuous functions on X with $\Lambda(1) = 1$. Then there exists a unique, regular, Borel probability measure μ on the natural Borel $\sigma-$algebra $\mathcal{U}$ of X such that μ represents Λ, that is,

$$\Lambda(f) = \int_X d\mu(x) f(x) \quad \forall\, f \in C(X) \tag{III.5.7}$$

Notice that regularity of μ on a compact Hausdorff space X reduces to the fact that the measure of every measurable set can be approximated arbitrarily well by open or compact (and hence closed since in a Hausdorff space every compact subset is closed, see [140]) sets respectively. Also, Lusin's theorem simplifies to the statement that every bounded measurable function can be approximated arbitarily well be continuous functions with the same bound up to sets of measure zero.

The notion of faithfulness actually comes from representation theory. Indeed, the origin of positive linear functionals in physics are usually states, that is, positive linear functionals ω on a unital C^*-algebra $\mathcal{A}$ (see section III.4), which is not necessarily Abelean like the C^*-algebra $C(X)$ for a compact Hausdorff space X, such that $\omega(1) = 1$. Here a positive linear functional is a map $\omega : \mathcal{A} \to \mathbb{C};\ a \mapsto \omega(a)$ which satisfies $\omega(a^* a) \geq 0$ for any $a \in \mathcal{A}$. Elements a of $\mathcal{A}$ of the form

b^*b are called positive, denoted $a \geq 0$ (equivalently, $a \geq 0$ iff for its spectrum $\sigma(a) \subset \mathbb{R}^+$ holds). One writes $a \geq a'$ if $a - a' \geq 0$ which equips $\mathcal{A}$ with a partial order. We will see in section III.6 that positive linear functionals give rise to a representation π of the algebra on a Hilbert space via the GNS construction. If the unital C^*-algebra is Abelean then we can always think of it as an algebra of continuous functions on a compact Hausdorff space via the Gel'fand isomorphism and if the associated measure is faithful, that is, the state is positive definite then the representation is faithful (or non-degenerate), that is, $\pi(f) = 0$ if and only if $f = 0$.

Notice that every positive linear functional ω on a unital C^* algebra $\mathcal{A}$ is automatically bounded (continuous):
If $||.||$ denotes the norm on $\mathcal{A}$ and * the involution then for any self-adjoint element $a = a^*$ we have $-||a||\cdot 1 \leq a \leq ||a||\cdot 1$ since $||a|| \geq r(a)$ (spectral radius). Hence $\omega(||a||\cdot 1 \pm a) = ||a||\omega(1) \pm \omega(a) \geq 0$. Since $\omega(1) \geq 0$ because $1 = 1^*1$ is positive, it follows that in particular $\omega(a) \in \mathbb{R}$ for self-adjoint a so that $|\omega(a)|/||a|| \leq \omega(1)$. If a is arbitrary we can decompose it uniquely into self-adjoint elements $a = a_+ + ia_-$ with $a_\pm = a_\pm^*$ and thus

$$4||a_\pm^2|| = ||(a^*)^2 + a^2 \pm (a^*a + aa^*)|| \leq ||(a^*)^2|| + ||a^2|| + ||a^*a|| + ||aa^*|| = 4||a||^2$$

where we have made use twice of the C^*-algebra property $||a^*a|| = ||a||^2$. It follows that

$$|\omega(a)|^2 = |\omega(a_+) + i\omega(a_-)|^2 = |\omega(a_+)|^2 + |\omega(a_-)|^2 \leq \omega(1)[||a_+||^2 + ||a_-||^2] \leq 2\omega(1)||a||^2$$

so a bound is given by $2\omega(1)$. One can actually show that a sharper bound is given by $\omega(1)$ even for unital Banach algebras with involution.

We now turn to another direction within measure theory.

Definition III.5.10
Let $(X, \mathcal{U}, \mu)$ be a measure space with a positive probability measure μ on X. Let $\lambda :\ G \times X \to X;\ (g, x) \mapsto \lambda_g(x)$ be a measure preserving group action, that is, $(\lambda_g)_*\mu := \mu \circ \lambda_g^{-1} = \mu$ for all $g \in G$, in particular, λ_g preserves $\mathcal{U}$. The group action is called ergodic if the only invariant sets, that is, sets $S \in \mathcal{U}$ with $\lambda_g(S) = S$ for all $g \in G$, have measure zero or one.

The definition captures exactly the intuitive idea of an ergodic group action, namely that it spreads any set all over X without changing its measure. It follows from the definition that a measure preserving group action induces a unitary transformation on $L_2(X, d\mu)$ by the pull-back, that is,

$$(\hat{U}(g)f)(x) := (\lambda_g^* f)(x) = f(\lambda_g(x)) \qquad \text{(III.5.8)}$$

Since the closed linear span of characteristic functions of measurable sets is all of $L_2(X, d\mu)$ as we have seen above, it follows that ergodicity is equivalent with the condition that $\hat{U}(g)f = f$ $\mu-$a.e. for all $g \in G$ implies that $f =$const. a.e. (Proof: If λ is ergodic and $f = \sum_k z_k \chi_{U_k}$ then $\hat{U}(g)f = \sum_k z_k \chi_{\lambda_{g^{-1}}(U_k)} = f$ a.e. for all $g \in G$ implies that all U_k must be invariant under λ, hence that all of them have measure zero or one. If U_k has measure zero then $\chi_{U_k} = 0$ a.e., if U_k has measure one then $X - U_k$ has measure zero so $\chi_{U_k} = \chi_X = 1$ a.e. The converse implication is similar).

Theorem III.5.5 (von Neumann Mean Ergodic Theorem)
Let $\mathbb{R} \to G;\ t \mapsto g_t$ be a one parameter group and $\hat{U} :\ G \to \mathcal{B}(L_2(X, d\mu)$ be a unitray representation of G. Let $\hat{P}$ be the projection on the closure of the set of a.e. invariant vectors under $\hat{U}(g_t)$, $t \in \mathbb{R}$. Then

$$(\hat{P}f)(x) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} dt (\hat{U}(g_t)f)(x) \ \ \mu - a.e. \qquad \text{(III.5.9)}$$

For a proof see for instance [129]. We conclude that λ restricted to $t \to g_t$ is ergodic if and only if

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt f(\lambda_{g_t}(x)) = [\int_X d\mu(x')f(x')] \cdot 1 \quad \mu - \text{a.e.} \tag{III.5.10}$$

Namely, if $t \to \lambda_{g_t}$ is ergodic, then the set of a.e. invariant vectors is given by the constant functions whence $\hat{P}f \propto 1$, that is,

$$\hat{P}f =< 1, \hat{P}f > \cdot 1 =< \hat{P}1, f > 1 =< 1, f > \cdot 1 = [\int_X d\mu(x)f(x)] \cdot 1 \tag{III.5.11}$$

since $1(x) = 1$ and the definition of the inner product. Comparing with $\hat{P}f$ from (III.5.9) gives the claimed result (III.5.10). Conversely, if (III.5.10) holds then the right hand side is constant almost everywhere and equals $\hat{P}f$ hence $t \to \lambda_{g_t}$ is ergodic by the above remark.

Criterion (III.5.10) is interesting for the following reason: Suppose that $\mu_1 \neq \mu_2$ are different measures on the same measurable space $(X, \mathcal{U})$, and that $t \to \lambda_{g_t}$ is a measure preserving, ergodic group action with respect to both of them. Then

$$[\int_X d\mu_1(x')f(x')] \cdot 1 =_{\mu_1-\text{a.e.}} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt f(\lambda_{g_t}(x)) =_{\mu_2-\text{a.e.}} [\int_X d\mu_2(x')f(x')] \cdot 1 \tag{III.5.12}$$

for any $f \in L_1(X, d\mu_1) \cap L_1(X, d\mu_2)$. Now the left and right hand side in (III.5.12) do not depend at all on the point x on which the middle term depends. Thus, if we can find $f \in L_1(X, d\mu_1) \cap L_1(X, d\mu_2) \neq \emptyset$ such that the constants $[\int_X d\mu_1(x')f(x') \neq [\int_X d\mu_2(x')f(x')]$ are different from each other then the middle term must equal the left hand side whenver $x \in \text{supp}(\mu_1)$ and it must equal the right hand side whenver $x \in \text{supp}(\mu_2)$. This is no contradiction iff μ_1, μ_2 are mutually singular with respect to each other. Hence ergodicity gives a simple tool for investigating the singularity structure of measures with respect to each other and one easily shows that Gaussian measures with different covariances (e.g. scalar fields with different masses) are built on mutually singular measures.

Definition III.5.11
A one parameter-group of measure preserving transformations $t \to \lambda_{g_t}$ is called mixing provided that

$$\lim_{t\to\infty} < f, \hat{U}(g_t)f' >=< f, 1 > < 1, f' > \tag{III.5.13}$$

It is easy to see that mixing implies ergodicity: Suppose that f' is invariant a.e. under the one-parameter group. Then by (III.5.13) and inserting the identity $1_{L_2} = \hat{P} \oplus [1_{L_2} - \hat{P}]$, where $\hat{P} = |1 >< 1|$ denotes the projection onto span($\{1\}$), gives

$$\begin{aligned} < f, f' > &= < f, 1 > < 1, f' > + < f, [1_{L_2} - \hat{P}]f' >=< f, 1 > < 1, f' > \\ &\Rightarrow < f, [1_{L_2} - \hat{P}]f' >= 0 \; \forall \; f \in L_2(X, d\mu) \end{aligned} \tag{III.5.14}$$

hence $[1_{L_2} - \hat{P}]f' = 0$ so that f' =const. a.e., that is, ergodicity.

III.6 Spectral Theorem and GNS-Construction

As an application of appendices III.4 and III.5 in addition to the general theory of the main text we present an elegant proof of the spectral theorem and sketch the GNS construction due to Gel'fand, Naimark and Segal.

Let $\mathcal{H}$ be a Hilbert space and a a bounded, linear, normal operator on $\mathcal{H}$, that is $||a|| = \sup_{\psi\neq 0}||a\psi||/||\psi|| < \infty$ where $||\psi||^2 =<\psi,\psi>$ denotes the Hilbert space norm and $[a,a^\dagger]=0$ where the bounded operator $a^\dagger$ is defined by $<a^\dagger\psi,\psi'>:=<\psi,a\psi'>$. More precisely, consider the linear form on $\mathcal{H}$ defined by

$$l_\psi : \mathcal{H} \to \mathbb{C};\psi' \to <\psi, a\psi'> \qquad \text{(III.6.1)}$$

This linear form is continuous since $|l_\psi(\psi')| \le ||\psi||\;||a||\;||\psi'||$ by the Schwarz inequality. Hence, by the Riesz lemma there exists $\xi_\psi \in \mathcal{H}$ such that $l_\psi =<\xi_\psi,.>$ and since l_ψ is conjugate linear in ψ it follows that $\psi \mapsto \xi_\psi := a^\dagger\psi$ actually defines a linear operator. Finally, $a^\dagger$ is bounded because

$$||a^\dagger\psi||^2 = |<\psi, aa^\dagger\psi>| \le ||\psi||\;||aa^\dagger\psi|| \le ||\psi||\;||a||\;||a^\dagger\psi|| \qquad \text{(III.6.2)}$$

again by the Schwarz inequality.

Let $\mathcal{A}$ be the unital, Abelean C^*-algebra generated by $1,a,a^\dagger$. It is Abelean since a is normal and the C^*-property follows from the following observation: Let $b\in\mathcal{A}$, then b is also normal and $||b\psi||^2 =<\psi,b^\dagger b\psi>= ||b^\dagger\psi||^2$ so that $||b|| = ||b^\dagger||$ for any $b\in\mathcal{A}$. Now by the Schwarz inequality $||b\psi||^2 = |<\psi,b^\dagger b\psi>| \le ||\psi||\;||b^\dagger b\psi||$ implying that $||b||^2 = ||b^\dagger||^2 \le ||b^\dagger b||$. On the other hand $||b^\dagger b|| \le ||b||\;||b^\dagger||$ due to submultiplicativity.

Consider the spectrum $\Delta(\mathcal{A}) = \text{Hom}(\mathcal{A},\mathbb{C})$ and the map $z:\ \Delta(\mathcal{A})\to\mathbb{C};\ \chi\mapsto\chi(a)$ which is continuous by the definition of the Gel'fand topology on the spectrum. We have seen already that the range of this map coincides with $\sigma(a)$. Moreover, z is injective because $\chi(a)=\chi'(a)$ implies that χ,χ' coincide on all polynomials of $a,a^\dagger$ since they are homomorphisms, and thus on allof $\mathcal{A}$ by continuity whence $\chi=\chi'$. Thus, z is a continuous bijection between the spectra of $\mathcal{A}$ and a respectively. Since a is bounded, both spectra are compact Hausdorff spaces. Now a continuous bijection between compact Hausdorff spaces is automatically a homeomorphism. (Proof: Let $f: X\to Y$ be a continuous bijection and let X be compact and Y Hausdorff. We must show that $f(U)$ is open in Y for every open subset $U\subset X$, or by taking complements, that images of closed sets are closed. Now since X is compact, it follows that every closed set U is also compact. Since f is continuous, it follows that $f(U)$ is compact. Since Y is Hausdorff it follows that $f(U)$ is closed. See theorems 5.3 and 5.5 of [140]). We conclude that we can identify $\Delta(\mathcal{A})$ topologically with $\sigma(a)$. By defintion the polynomials p in $a,a^\dagger$ lie dense in $\mathcal{A}$ and we have for $\chi\in\Delta(\mathcal{A})$ that

$$\chi(p(a,a^\dagger)) = p(\chi(a),\overline{\chi(a)}) = p(z(\chi),\overline{z(\chi)}) = [p\circ(z,\bar{z})](\chi) = p(a,a^\dagger)^\vee(\chi) \qquad \text{(III.6.3)}$$

so that the Gel'fand isometric isomorphism can be thought of as a map $\vee:\ \mathcal{A}\to C(\sigma(a));\ b\mapsto\check{b}$ with $\check{b}(z) = \chi(b)_{z=\chi(a)}$.

Now consider any state $\psi\in\mathcal{H}$ with $||\psi||=1$. Then

$$\omega_\psi:\ \mathcal{A}\to\mathbb{C};\ b\mapsto<\psi,b\psi> \qquad \text{(III.6.4)}$$

is obviously a state on $\mathcal{A}$. Via the Gel'fand transform we obtain a positive linear functional on $C(\sigma(a))$ by

$$\Lambda_\psi:\ C(\sigma(a))\to\mathbb{C};\ \check{b}\mapsto\omega_\psi(b) \qquad \text{(III.6.5)}$$

and since $\sigma(a)$ is a compact Hausdorff space we can apply the Riesz representation theorem in order to find a unique, regular Borel measure μ_ψ on $\sigma(a)$ such that

$$\omega_\psi(b) = \int_{\sigma(a)} d\mu_\psi(z)\check{b}(z) \tag{III.6.6}$$

The measure μ_ψ is caled a spectral measure. The meaning of this formula is explained by the following definition.

Definition III.6.1

i)

A representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ of a C^*-algebra is a $^*-$homomorphism into the $^*-$algebra of bounded operators on a Hilbert space. The representation is said to be faithful if $a \neq 0$ implies $\pi(a) \neq 0$. Two representations $\pi_I; \mathcal{A} \to \mathcal{B}(\mathcal{H}_I)$; $I = 1, 2$ are called equivalent iff there exists a Hilbert space isomorphism $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $\pi_2(a) = U\pi_1(a)U^{-1}$ for all $a \in \mathcal{A}$. Finally, a representation is called non-degenerate if $ker(\pi) := \{\psi \in \mathcal{H};\ \pi(a)\psi = 0\ \forall\ a \in \mathcal{A}\}$ is given by $\{0\}$.

ii)

Let ω be a state on a unital $C^* -$ algebra and define the null space $\mathcal{N}_\omega := \{a \in \mathcal{A};\ \omega(a^*a) = 0\}$. The GNS representation with respect to ω

$$\pi_\omega :\ \mathcal{A} \to \mathcal{B}(\mathcal{H}_\omega) \text{ where } \mathcal{H}_\omega := \overline{\mathcal{A}/\mathcal{N}_\omega} := \overline{\{[a];\ a \in \mathcal{A}\}} \tag{III.6.7}$$

where the overbar denotes completion and $[.] :\ \mathcal{A} \to \mathcal{A}/\mathcal{N}_\omega;\ a \mapsto [a] := \{a + b;\ b \in \mathcal{N}_\omega\}$ is the quotient map, is densely defined by

$$\pi_\omega(a)[b] := [ab] \tag{III.6.8}$$

and extended by continuity. The Hilbert space $\mathcal{H}_\omega$ is equipped with the inner product

$$< [a], [b] >_{\mathcal{H}_\omega} := \omega(a^*b) \tag{III.6.9}$$

The Hilbert space state $\Omega_\omega := [1]$ is cyclic for $\mathcal{H}_\omega$, that is, a dense set of Hilbert space space states is obtained as $\{\pi_\omega(a)\Omega_\omega;\ a \in \mathcal{A}\}$. Moreover,

$$\omega(a) =< \Omega_\omega, \pi_\omega(a)\Omega_\omega >_{\mathcal{H}_\omega} \tag{III.6.10}$$

To see that this definition makes sense, one notices that $\mathcal{N}_\omega$ is a closed left ideal in $\mathcal{A}$ so that (III.6.8), (III.6.9) are well-defined (using that the right hand side of (III.6.9) defines a positive semidefinite sesquilinear form on $\mathcal{H} := \mathcal{A}$ and hence the Schwarz inequality applies) and that $||\pi_\omega(a)|| = ||a||$ is indeed bounded. One can show that the triple $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ is fixed by condition (III.6.10) up to unitary equivalence. Notice that the state is not required to be pure (i.e. cannot be written as a convex linear combination of other states) but if it is then one can show that the representation is irreducible (does not contain invariant subspaces different from itself and $\{0\}$).

It is almost clear that every non-degenerate representation is an orthogonal sum of cyclic representations: Take an arbitrary element $0 \neq \psi \in \mathcal{H}$ and construct $\mathcal{H}_\psi := \overline{\{\pi(a)\psi;\ a \in \mathcal{A}\}}$. If $\mathcal{H}_\psi^\perp \neq \{0\}$ take $\psi' \in \mathcal{H}_\psi^\perp$ and iterate. The rigrous proof makes use of the axiom of choice and will be left to the reader.

Coming back to our concrete C^*-algebra $\mathcal{A}$ generated by a normal, bounded operator $a \in \mathcal{B}((\mathcal{H})$ on a given Hilbert space $\mathcal{H}$ we see that it is represented as $\pi(b) = b$ on $\mathcal{H}$ and that this representation is non-degenerate because $\mathcal{A}$ contains the identity operator. We then find an index set A, vectors ψ_α and closed, mutually orthogonal subspaces $\mathcal{H}_\alpha := \overline{\{b\psi_\alpha;\ b \in \mathcal{A}\}}$ containing ψ_α such that $\mathcal{H} = \oplus_{\alpha\in A}\mathcal{H}_\alpha$.

By construction, the subspaces $\mathcal{H}_\alpha$ are invariant for $\mathcal{A}$. Then any vector $\psi \in \mathcal{H}$ is (in the closure of vectors of) the form $\psi = \sum_{\alpha\in A} b_\alpha \psi_\alpha$ with $b_\alpha \in \mathcal{A}$ and we have

$$< \psi, \psi' >= \sum_{\alpha\in A} < \psi_\alpha, b_\alpha^\dagger b'_\alpha \psi_\alpha > \tag{III.6.11}$$

Using the result (III.6.6) we may write this as

$$< \psi, \psi' >= \sum_{\alpha\in A} \int_{\sigma(a)} d\mu_{\psi_\alpha}(z) \overline{\check{b}_\alpha(z)} \check{b}'_\alpha(z) \tag{III.6.12}$$

where we have used that $(b^\dagger b')^\vee = \bar{\check{b}}\check{b}'$. This formula suggests to introduce the Hilbert spaces $L_2(\sigma(a), d\mu_{\psi_\alpha})$ as well as the space $\sigma := \bigcup_{\alpha\in A} \sigma(a)_\alpha$ (disjoint union of copies of $\sigma(a)$) and a measure μ on it defined by $\mu_{|\sigma(a)_\alpha} := \mu_{\psi_\alpha}$. Notice that measurable sets are of the form $\bigcup_{\alpha\in B\subset A} U_\alpha$ where U_α is measurable in $\sigma(a)_\alpha$, B can be any subindex set and that unions, intersections and differences of measurable sets are performed componentwise. Let us now define the Hilbert space $L_2(\sigma, d\mu)$. An element $\check{\psi}$ of $L_2(\sigma, d\mu)$ is a square integrable function on σ with respect to the measure μ and may be defined in terms of an array of functions $\check{\psi}_\alpha \in L_2(\sigma(a)_\alpha, d\mu_{\psi\alpha)}$ through $\check{\psi}_{|\sigma(a)_\alpha} = \check{\psi}_\alpha$. Notice that indeed

$$\begin{aligned}
< \check{\psi}, \check{\psi}' >_{L_2(\sigma,d\mu)} &= \int_\sigma d\mu(z) \overline{\check{\psi}(z)} \check{\psi}'(z) \\
&= \sum_{\alpha\in A} \int_{\sigma(a)_\alpha} d\mu(z) \overline{\check{\psi}(z)} \check{\psi}'(z) = \sum_{\alpha\in A} \int_{\sigma(a)_\alpha} d\mu_{|\sigma(a)_\alpha}(z) [\overline{\check{\psi}(z)} \check{\psi}'(z)]_{|\sigma(a)_\alpha} \\
&= \sum_{\alpha\in A} \int_{\sigma(a)} d\mu_{\psi_\alpha}(z) \overline{\check{\psi}_\alpha(z)} \check{\psi}'_\alpha(z)
\end{aligned} \tag{III.6.13}$$

explaining the requirement that $\check{\psi}_\alpha \in L_2(\sigma(a)_\alpha, d\mu_{\psi_\alpha})$. Here we have made use of $\sigma-$additivity, that is, $\mu(\bigcup_\alpha U_\alpha) = \sum_\alpha \mu(U_\alpha) = \sum_\alpha \mu_{\psi_\alpha}(U_\alpha)$ for the mutually disjoint sets $U_\alpha \subset U$. Comparing (III.6.12) and (III.6.13) we see that we can identify $L_2(\sigma, d\mu)$ with $\oplus_{\alpha\in A} L_2(\sigma(a)_\alpha, d\mu_{\psi_\alpha})$ and obtain a unitary transformation

$$U:\ \mathcal{H} \to L_2(\sigma, d\mu);\ \ \psi = \sum_{\alpha\in A} b_\alpha \psi_\alpha \mapsto \check{\psi} \text{ where } \check{\psi}_{|\sigma(a)_\alpha} := \check{b}_\alpha \tag{III.6.14}$$

Moreover, we have

$$U b\psi = U \sum_\alpha b b_\alpha \psi_\alpha = \check{\psi}' \text{ where } \check{\psi}'_{|\sigma(a)_\alpha} = (b b_\alpha)^\vee = \check{b}\check{b}_\alpha \tag{III.6.15}$$

which means that on each supspace $L_(\sigma(a)_\alpha, d\mu_{\psi_\alpha})$ the operator b is represented by multiplication by $\check{b}(z)$. In particular, if $b = a$ or $b = a^\dagger$ it is represented by multiplication by z or $\bar{z}$ since $\chi(a) = z$.

This simple corollary from Gel'fand spectral theory and the Riesz representation theorem is the spectral theorem for bounded operators. It obviously generalizes to the case that we have a family (a_I) of bounded operators which together with their adjoints mutually commute with each other. The only difference is that we now get a homeomorphism between $\Delta(\mathcal{A})$ and the joint spectrum $\prod_I \sigma(a_I)$ via $\chi \mapsto (\chi(a_I))_I$. We can also strip off the concrete Hilbert space context by considering an abstract unital C^*-algebra $\mathcal{A}$ where instead of vector states ψ_α we use states ω_α on $\mathcal{A}$ and apply the GNS construction. That for given $a \in \mathcal{A}$ there is always a state ω with $\omega(a^*a) > 0$ follows from the Hahn-Banach theorem applied to the vector space $X := \mathcal{A}$ and its one-dimesional supspace $Y := \text{span}(a^*a)$ with the bounding function appearing in the theorem given by the norm on X and by defining $\omega(a^*a) := ||a||^2$: The Hahn-Banach theorem guarantees that then ω can be extended as a positive linear functional to all of $\mathcal{A}$.

Theorem III.6.1
Let (a_I) be a self-adjoint collection of mutually commuting elements of a C^*-algebra $\mathcal{C}$. Then there exists a representation of the sub$-C^*-$algebra $\mathcal{A}$ generated by this collection on a Hilbert space $\mathcal{H}$ such that the $\pi(a_I)$ become multiplication operators.

The extension of the spectral theorem to unbounded self-adjoint operators operators on a Hilbert space can be traced back to the bounded case by using the following trick. (Recall that a densely defined operator a with domain $D(A)$ is called self-adjoint if $a^\dagger = a$ and $D(a^\dagger) = D(a)$ where

$$D(a^\dagger) := \{\psi \in \mathcal{H};\ \sup_{0 \neq \psi' \in D(a)} | < \psi, a\psi' > |/||\psi'|| < \infty\}$$

and $a^\dagger$ is uniquely defined on $\psi \in D(a^\dagger)$ via $< a^\dagger\psi, \psi' >=< \psi, a\psi' >$ for all $\psi \in D(a)$ through the Riesz lemma):
The spectrum of a will be an unbounded subset of the real line. Let f be a bijection $\mathbb{R} \to K$ where K is a compact one-dimensional subset of $\mathbb{C}$ and suppose that $f(a)$ is a bounded operator. Then we can apply the spectral theorem for bounded normal operators to $f(a)$ which then becomes a multiplication operator and if f^{-1} is a measurable function with respect to the spectral measure μ then also a itself is a multiplication operator. A popular tool is the Caley transform $a \to u := (a - i)(a + i)^{-1}$ which maps a to a unitary operator.

Finally, let us mention the spectral resolution. Let a be a bounded self-adjoint operator then by the spectral theorem there is a measure μ and a representation such that $< \psi, f(a)\psi >= \int_{\sigma(a)} d\mu_\psi(z) f(z)$ for any measurable function f and μ_ψ is the spectral measure of ψ in a cyclic representation. Let $S \subset \mathbb{R}$ be measurable and and consider the operators $p_S := \chi_S(a)$ called the spectral projections where χ_S is the characteristic function of S. Then $< \psi, p_S\psi >= \int_{\sigma(a)} d\mu_\psi(z)\chi_S(z)$. Let $p_z := \chi_{(-\infty,z)}(a)$ for $z \in \mathbb{R}$ then we see that we obtain the so-called projection valued measures

$$< \psi, dp_z\psi >= d < \psi, dp_z\psi >= d\mu_\psi(z) \tag{III.6.16}$$

whence

$$< \psi, f(a)\psi >= \int_{\mathbb{R}} < \psi, dp_z\psi > f(z) \tag{III.6.17}$$

for all $\psi \in \mathcal{H}$ or by the the polarization identity

$$f(a) = \int_{\mathbb{R}} dp_z f(z) \tag{III.6.18}$$

which is called the spectral resolution of $f(a)$.

III.7 Introduction to Refined Algebraic Quantization (RAQ)

RAQ provides strong guidelines of how to solve a given family of quantum constraints but unfortunately it is not an algorithm that one just has to apply in order to arrive at a satisfactory end result. In particular, as presently formulated it has its limitations since it does not cover the case that the constraints form an open algebra with structure functions rather than structure constants as it would be the case for a Lie algebra. Unfortunately, quantum gravity belongs to the open algebra category of constrained systems. We mainly follow Giulini and Marolf in [56].

Let $\mathcal{H}_{kin}$ be a Hilbert space, referred to as the kinematical Hilbert space because it is supposed to implement the adjointness – and canonical commutation relations of the elementary kinematical degrees of freedom. However, these degrees of freedom are not observables (classically they do not have vanishing Poisson brackets with the constraints on the constraint surface) and the Hilbert space is not the physical one on which the constraint operators would equal the zero operators. The role of $\mathcal{H}_{kin}$ is to give the constraint operators $(\hat{C}_I)_{I\in\mathcal{I}}$ a home, that is, there is a common dense domain $\mathcal{D}_{kin} \subset \mathcal{H}_{kin}$ which is supposed to be invariant under all the $\hat{C}_I$ and we also require that the $\hat{C}_I$ be closed operators (i.e. their adjoint is densely defined as well). We do not require them to be bounded operators. The label set $\mathcal{I}$ is rather arbitrary and usually is a combination of direct products of finite and infinite sets (e.g. tensor or gauge group indices times indices taking values in a separable space of smearing functions).

We will further require that the constraints form a first class system and that they actually form a Lie algebra, that is, there exist complex valued structure constants $f_{IJ}\,^K$ such that

$$[\hat{C}_I, \hat{C}_J] = f_{IJ}\,^K \hat{C}_K \tag{III.7.1}$$

where the summation over K performed here will involve an integral for generic $\mathcal{I}$. Notice that (III.7.1) makes sense due to our requirement on $\mathcal{D}_{kin}$. The case of an open algebra would correspond to the fact that the structure constants become operator valued as well and then it becomes an issue how to choose the operator ordering in (III.7.1), in particular, if constraint operators and structure constant operators are chosen to be self-adjoint and anti-self-adjoint respectively (which would be natural if their classical counterparts are classically real and imaginary valued respectively) then one would have to order (III.7.1) symmetrically which would be a desaster for solving the constraints, see below, which is why in the open case the constraints should not be chosen to be self-adjoint operators. Notice that there is no contradiction because self-adjointness usually is required to ensure that the spectrum (measurement values) of the operator lies in the real line, however, for constraint operators this requirement is void since we are only interested in their kernel and the only requirement is that the point zero belongs to the spectrum at all.

In order to allow for non-self-adjoint constraints, in what follows we will assume that the set $\mathcal{C} := \{\hat{C}_I;\ I \in \mathcal{I}\}$ is self-adjoint (i.e. contains with $\hat{C}_I$ also $\hat{C}_I^\dagger = \hat{C}_J$ for some J) which means that the dense domain $\mathcal{D}_{kin}$ is also a dense domain for the adjoints so that the constraints are explicitly closed operators. Let us now consider the self-adjoint set of kinematical observables $\mathcal{O}_{kin}$, that is, all operators on $\mathcal{H}_{kin}$ which have $\mathcal{D}_{kin}$ as common dense domain together with their adjoints. Obviously, $\mathcal{O}_{kin}$ contains $\mathcal{C}$. Consider the commutant of $\mathcal{C}$ within $\mathcal{O}_{kin}$, that is,

$$\mathcal{C}' := \{O \in \mathcal{O}_{kin};\ [C, O] = 0\ \forall\ C \in \mathcal{C}\} \tag{III.7.2}$$

It is clear that $\mathcal{C}'$ is a $^*-$subalgebra of $\mathcal{O}_{kin}$ since $[O^\dagger, \mathcal{C}] = -([O, \mathcal{C}])^\dagger = 0$ and $[OO', \mathcal{C}] = O[O', \mathcal{C}] + [O, \mathcal{C}]O' = 0$ for any $O, O' \in \mathcal{C}'$ since $\mathcal{C}^\dagger = \mathcal{C}$ is a self-adjoint set. Moreover, $\mathcal{C}$ might have a non-trivial center

$$\mathcal{Z} = \mathcal{C} \cap \mathcal{C}' \tag{III.7.3}$$

which generates a two-sided ideal $I_{\mathcal{Z}}$ in $\mathcal{C}'$ corresponding to classical functions that vanish on the constraint surface and is therfore physically uninteresting. Hence we will define the algebra of physical observables to be the quotient algebra

$$\mathcal{O}_{phys} := \mathcal{C}'/\mathcal{Z} \tag{III.7.4}$$

Usually the space $\mathcal{D}_{kin}$ comes with its own topology τ, different from the subspace topology inherited from the Hilbert space topology $||.||$ on $\mathcal{H}_{kin}$, generically a nuclear topology [59] so that $\mathcal{D}_{kin}$ becomes a Fréchet space (a space whose topology is generated by a countable family of semi-norms that separates the points of $\mathcal{D}_{kin}$ and such that $\mathcal{D}_{kin}$ is complete in the associated norm; a general locally convex topological vector space is not necessarily complete and the family of semi-norms need not to be countable (it is then not metrizable)). The intrinsic topology τ is then finer than $||.||$ since $\mathcal{D}_{kin}$ is complete but also dense in $\mathcal{H}_{kin}$ (if it would be coarser then a Cauchy sequence in $\mathcal{D}_{kin}$ with respect to the intrinsic topology would also be one in the Hilbert space topology and since $\mathcal{D}_{kin}$ is dense this completion would coincide with $\mathcal{H}_{kin}$). It follows that the space of continuous linear functionals $\mathcal{D}'_{kin}$ (with respect to the topology on $\mathcal{D}_{kin}$ contains $\mathcal{H}_{kin}$ since a Hilbert space is reflexive, that is, $\mathcal{H}'_{kin} = \mathcal{H}_{kin}$ by the Riesz lemma so the elements of $\mathcal{H}_{kin}$ are in particular continuous linear functionals on $\mathcal{D}_{kin}$ with respect to $||.||$ so that they are also continuous with respect to τ (a function stays continuous if one strengthens the topology on the domain space). Let (l^α) be a net in $\mathcal{H}'_{kin}$ converging to l then

$$\begin{aligned} ||l^\alpha - l||_{\mathcal{D}'_{kin}} &= \sup_{f\in\mathcal{D}_{kin}} \frac{|<l_\alpha - l, f>|}{||f||_{\mathcal{D}_{kin}}} = \sup_{f\in\mathcal{D}_{kin}} \frac{||f||_{\mathcal{H}_{kin}}}{||f||_{calD_{kin}}} \frac{|<l_\alpha - l, f>|}{||f||_{\mathcal{H}_{kin}}} \\ &\leq \sup_{f\in\mathcal{D}_{kin}} \frac{|<l_\alpha - l, f>|}{||f||_{\mathcal{H}_{kin}}} \leq \sup_{f\in\mathcal{H}_{kin}} \frac{|<l_\alpha - l, f>|}{||f||_{\mathcal{H}_{kin}}} = ||l^\alpha - l||_{\mathcal{H}'_{kin}} \end{aligned} \tag{III.7.5}$$

where we used $||f||_{\mathcal{H}_{kin}}/||f||_{calD_{kin}} \geq 1$. Thus it converges in $\mathcal{D}'_{kin}$ as well, that is, the topology on $\mathcal{D}'_{kin}$ is weaker than that of $\mathcal{H}_{kin}$. We thus have topological inclusions

$$\mathcal{D}_{kin} \hookrightarrow \mathcal{H}_{kin} \hookrightarrow \mathcal{D}'_{kin} \tag{III.7.6}$$

sometimes called a Gel'fand triple.

Unfortunately the definition of a Gel'fand triple requires a further input, the nuclear topology intrinsic to $\mathcal{D}_{kin}$ which we want to avoid since there seems no physical guiding principle (although then there are rather strong theorems available concerning the completeness of generalized eigenvectors [59]). We thus equip $\mathcal{D}_{kin}$ simply with the relative topology induced from $\mathcal{H}_{kin}$. The requirement that $\mathcal{D}_{kin}$ is dense is then no loss of generality since we may simply replace $\mathcal{H}_{kin}$ by the completion of $\mathcal{D}_{kin}$. Instead of the topological dual (which would coincide with $\mathcal{H}_{kin}$ we consider the algebraic dual $\mathcal{D}^*_{kin}$ of all linear functionals on $\mathcal{D}_{kin}$. This space is naturally equipped with the weak * topology of pointwise convergence, i.e. a net (l^α) converges to l iff the net of complex numbers $(l^\alpha(f))$ converges to $l(f)$ for any $f \in \mathcal{D}_{kin}$ (but not uniformly). Again we can consider $\mathcal{H}_{kin}$ as a subspace of $\mathcal{D}^*_{kin}$ and since a net converging in norm certainly converges pointwise we have again topological inclusions

$$\mathcal{D}_{kin} \hookrightarrow \mathcal{H}_{kin} \hookrightarrow \mathcal{D}^*_{kin} \tag{III.7.7}$$

which in abuse of notation we will still refer to as Gel'fand triple. Thus, the only input left is the choice of $\mathcal{D}_{kin}$ for which, however, there are no general selection principles available at the moment (see however [56] for further discussion).

The reason for blowing up the structure beyond $\mathcal{H}_{kin}$ is that generically the point zero does not lie in the discrete part of the spectrum of $\mathcal{C}$, that is, if we look for solutions to the constraints in the form $\hat{C}_I\psi = 0$ for all $I \in \mathcal{I}$ for $\psi \in \mathcal{H}_{kin}$, then there are generically not enough solutions because ψ would be an eigenvector with eigenvalue zero but since zero does not lie in the discrete spectrum the eigenvectors do not form the entire solution space. This is precisely what happens with the diffeomorphism constraint for the case of quantum gravity where the only eigenvectors are the constant functions. We therefore look for generalized eigenvectors $l \in \mathcal{D}_{kin}^*$ in the algebraic dual for which we require

$$[(\hat{C}_I^\dagger)'l](f) := l(\hat{C}_I f) = 0 \; \forall \; I \in \mathcal{I}, \; f \in \mathcal{D}_{kin} \tag{III.7.8}$$

where the dual action of an operator $\hat{O} \in \mathcal{O}_{kin}$ on $l \in \mathcal{D}_{kin}^*$ is defined by

$$[\hat{O}'l](f) := l(\hat{O}^\dagger f) \; \forall \; f \in \mathcal{D}_{kin} \tag{III.7.9}$$

Notice that since we required $\mathcal{C}$ to be a self-adjoint can avoid taking the adjoint in (III.7.8) by passing to self-adjoint representatives $\hat{C}_I$. Due to the adjoint operation in (III.7.9) we have an anti-linear representation of $\mathcal{O}_{kin}$ on $\mathcal{D}_{kin}^*$ which descends to an anti-linear representation of $\mathcal{O}_{phys}$ on the space of solutions $\mathcal{D}_{phys}^* \subset \mathcal{D}_{kin}^*$ to (III.7.8).

At this point, the space $\mathcal{D}_{phys}^*$ is just a subspace of $\mathcal{D}_{kin}^*$. Ww would like to equip a subspace H_{phys} of it with a Hilbert space topology. The reason for not turning all of $\mathcal{D}_{phys}^*$ into $\mathcal{H}_{phys}$ is that then $\mathcal{O}_{phys}$ would be realized as an algebra of bounded operators on $\mathcal{H}_{phys}$ since they are defined everywhere on $\mathcal{D}_{phys}^*$ which would be unnatural if the corresponding classical functions are unbounded. In particular, the topology on $\mathcal{H}_{phys}$, as a complete norm topology, should be finer than the relative topology induced from $\mathcal{D}_{kin}^*$. The idea is then to consider $\mathcal{D}_{phys}^*$ as the algebraic dual of a dense subspace $\mathcal{D}_{phys} \subset \mathcal{H}_{phys}$ so that all of $\mathcal{O}_{phys}$ is densely defined there. In other words we get a second Gel'fand triple

$$\mathcal{D}_{phys} \hookrightarrow \mathcal{H}_{phys} \hookrightarrow \mathcal{D}_{phys}^* \tag{III.7.10}$$

with an anti-linear representation of $\mathcal{O}_{phys}$ on $\mathcal{H}_{phys}$ defined by (III.7.9).

The choice of the inner product on $\mathcal{H}_{phys}$ is guided by the requirement that the adjoint in the physical inner product, denoted by $\star$, represents the adjoint in the kinematical one, that is,

$$<\psi, \hat{O}'\psi'>_{phys} = <(\hat{O}')^\star\psi, \psi'>_{phys} = <(\hat{O}^\dagger)'\psi, \psi'>_{phys} \tag{III.7.11}$$

for all $\psi, \psi' \in \mathcal{D}_{phys}$. The canonical commutation relations among observables are automatically implemented because by construction $\mathcal{H}_{phys}$ carries a representation of $\mathcal{O}_{phys}$ on which the correct algebraic relations were already implemented as an abstract algebra.

A systematic construction of the physical inner product is available if we have an anti-linear (so-called) rigging map

$$\eta : \; \mathcal{D}_{kin} \to \mathcal{D}_{phys}^*; \; f \mapsto \eta(f) \tag{III.7.12}$$

at our disposal which must be such that
1) the following is a positive semi-definite sesquilinear form (linear in f, anti-linear in f')

$$<\eta(f), \eta(f')>_{phys} := [\eta(f')](f) \; \forall \; f, f' \in \mathcal{D}_{kin} \tag{III.7.13}$$

2) For any $\hat{O} \in \mathcal{O}_{phys}$ we have

$$\hat{O}'\eta(f) = \eta(\hat{O}f) \; \forall \; f \in \mathcal{D}_{kin} \tag{III.7.14}$$

which makes sure that the dual action preserves the space of solutions since $\hat{C}'\hat{O}'\eta(f) = 0$. Notice that bot the left and the right hand side in (III.7.14) are antilinear in $\hat{O}$.

We could then define $\mathcal{D}_{phys} := \eta(\mathcal{D}_{kin})/\ker(\eta)$ (with the kernel being understood with respect to $||.||_{phys}$) and complete it with respect to (III.7.13) to obtain $\mathcal{H}_{phys}$. Notice that (III.7.11) is satisfied because for $\psi = \eta(f), \psi' = \eta(f')$ we have

$$\begin{aligned} <\psi, \hat{O}'\psi'>_{phys} &= <\eta(f), \eta(\hat{O}f')>_{phys} = [\eta(\hat{O}f')](f) = [\hat{O}'\eta(f')](f) = \eta(f')(\hat{O}^\dagger f) \\ &= <\eta(\hat{O}^\dagger f), \eta(f')>_{phys} = <(\hat{O}^\dagger)'\psi, \psi'>_{phys} \end{aligned} \tag{III.7.15}$$

To see that $\mathcal{H}_{phys}$ is a subspace of $\mathcal{D}^*_{phys}$ with a finer topology, notice that the map $J:\ \mathcal{H}_{phys} \to \mathcal{D}^*_{phys}$ defined by $[J(\psi)](f) := <\psi, \eta(f)>_{phys}$ is an injection because $J(\psi)$ vanishes iff ψ is orthogonal to all $\eta(f)$ with respect to $<.,.>_{phys}$ which means that $\psi = 0$ because the image of η is dense. Hence J is an embedding (injective inclusion) of linear spaces. Moreover, J is evidently continuous: if $||\psi^\alpha - \psi||_{phys} \to 0$ then $J(\psi^\alpha) \to J(\psi)$ in the weak * topology iff $[J(\psi^\alpha)](f) \to [J(\psi)](f)$ for any $f \in \mathcal{D}_{kin}$ which is clearly the case. So convergence in $\mathcal{H}_{phys}$ implies convergence of $J(\mathcal{H}_{phys})$, hence the Hilbert space topology is stronger than the relative topology on $J(\mathcal{H}_{phys})$.

Thus, the existence of a rigging map solves the problem of defining a suitable inner product. A heuristic idea of how to construct η is through the group averaging proposal: Since $\mathcal{C}$ is a self-adjoint set we may assume w.l.g that the $\hat{C}_I$ are self-adjoint, and since they form a Lie algebra we may in principle exponentiate this Lie algebra (using the spectral theorem) and obtain a group of operators $t^I \to \exp(t^I \hat{C}_I)$ where $t^I \in T$ is some set depending on the constraints. Let then

$$\eta(f) := \overline{\int_T d\mu(t) \exp(t^I \hat{C}_I) f} \tag{III.7.16}$$

with a translation invariant measure μ on T. One easily sees that with

$$[\eta(f)](f') := \int_T d\mu(t) <\exp(t^I \hat{C}_I) f, f'>_{kin} \tag{III.7.17}$$

formally $[\eta(f)](\hat{C}_I f') = 0$. Of course, one must check case by case whether T, μ exist and that η has the required properties.

Let us make some short comments about the open algebra case:
Suppose that the classical constraint functions C_I and the structure functions $f_{IJ}\,^K$ are real and imaginary valued respectively. As mentioned already, it is now excluded to choose the corresponding operators to be (anti)-self-adjoint opertors since this would require the ordering

$$[\hat{C}_I, \hat{C}_J] = \frac{1}{2}(\hat{f}_{IJ}\,^K \hat{C}_K + \hat{C}_K \hat{f}_{IJ}\,^K) \tag{III.7.18}$$

and would lead to the following quantum anomaly: If we impose the condition (III.7.8) then we would find for an element $l \in \mathcal{D}^*_{phys}$ that

$$((\hat{f}_{IJ}\,^K)' \hat{C}'_K + \hat{C}'_K (\hat{f}_{IJ}\,^K)') l = [\hat{C}'_K, (\hat{f}_{IJ}\,^K)'] l = 0 \tag{III.7.19}$$

which means that l is not only annihilated by the dual constraint operators but also by (III.7.19) which is not necessarily proportional to a dual constraint operator any longer, implying that the physical Hilbert space will be too small. If on the other hand we do not choose the $\hat{C}_I$ to be self-adjoint, the anomaly problem is potentially absent but now it is no longer true that $[\hat{C}'_I l](f) = l(\hat{C}_I f)$, in other words, the question arises whether it is $\hat{C}'_I l = 0$ or $(\hat{C}^\dagger_I)' l = 0$ that we should impose ? The answer is that this just corresponds to a choice of operator ordering since the classical limit of both $\hat{C}_I$ and $\hat{C}^\dagger_I$ is given by the real valued function C_I and thus the answer is that the correct ordering is the one in which the algebra is, besides being densely defined and closed, also free of anomalies. Thus, in the open algebra case we may proceed just as above with the additional requirement of anomaly freeness. Of course, group averaging does not work since we cannot eponentiate the algebra any longer.

We conclude this section with an example in order to illustrate the procedure:
Suppose $\mathcal{H}_{kin} = L_2(\mathbb{R}^2, d^2x)$ and $\hat{C} = \hat{p}_1 = -i\partial/\partial x_1$. Obviously the kinematical Hilbert space implements the adjointness and canonical commutation relations among the basic variables x_1, x_2, p_1, p_2. A nuclear space choice would be $\mathcal{D}_{kin} = \mathcal{S}(\mathbb{R}^2)$ (test functions of rapid decrease). The functions l annihilated by $\hat{C}$ are those that do not depend on x^1 and are thus not normalizable. However, we can define them as elements of $\mathcal{D}^*_{kin}$ by $l(f) :=< l, f >_{kin}= \int_{\mathbb{R}^2} d^2x\overline{l(x)}f(x)$ which converges pointwise. Clearly $l(\hat{C}f) = 0$ if $l_{,x^1} = 0$. The physical observable algebra consists of operators not involving $\hat{x}_1$ and after taking the quotient with respect to the constraint ideal they involve only $\hat{p}_2, \hat{x}_2$. Obviously they leave the space $\mathcal{D}^*_{phys}$ invariant, consisting of those elements of $\mathcal{D}^*_{kin}$ that are x^1-independent. The physical Hilbert space that suggests itself (implementing the correct reality condition) is therefore $\mathcal{H}_{phys} = L_2(\mathbb{R}, dx_2)$ which is a proper subspace of $\mathcal{D}^*_{phys}$ and we have $D_{phys} = \mathcal{S}(\mathbb{R})$. Now an appropriate rigging map is obtained indeed by

$$\eta(f)(x_1, x_2) := \overline{\int_{\mathbb{R}} dt \exp(it\hat{p}_1)f(x_1, x_2)} = \overline{\int_{\mathbb{R}} dx_1 f(x_1, x_22)} = 2\pi\overline{\delta(\hat{C})f(x_1, x_2)}$$

since $\hat{p}_1$ generates x_1 translations, produces functions independent of x^1 and dt is an invariant measure on $T = \mathbb{R}$. Notice that the integral converges because f is of rapid decrease. Notice also that we could define the delta distribution of the constraint, using the spectral theorem. In the case of an Abelean self-adjoint constraint algebra a reasonable ansatz for a rigging map is always given by

$$\eta(f) = \overline{\prod_{I\in\mathcal{I}} \delta(\hat{C}_I)f} \tag{III.7.20}$$

We have

$$\begin{aligned} < \eta(f), \eta(f') >_{phys} &:= \eta(f')[\eta(f)] = \int_{\mathbb{R}} dt \int_{\mathbb{R}^2} d^2x\overline{f'(x_1 + t, x_2)}f(x_1, x_2) \\ &= \int_{\mathbb{R}} dx_2 \overline{[\int dx_1' f'(x_1', x_2)]}[\int dx_1 f(x_1, x_2)] \end{aligned} \tag{III.7.21}$$

which is the same inner product as chosen above.

Bibliography

[1] C. Rovelli, "Loop Quantum Gravity", Review written for electronic journal "Living Reviews", gr-qc/9710008
C. Rovelli, "Strings, Loops and Others : a critical Survey of the present Approaches to Quantum Gravity", plenary lecture given at 15th Intl. Conf. on Gen. Rel. and Gravitation (GR15), Pune, India, Dec 16-21, 1997, gr-qc/9803024
M. Gaul, C. Rovelli, "Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance", Lect.Notes Phys.541:277-324,2000, [gr-qc/9910079]

[2] G. Horowitz, "Quantum Gravity at the Turn of the Millenium", gr-qc/0011089
C. Rovelli, "Notes for a Brief History of Quantum Gravity", gr-qc/0006061
S. Carlip, "Quantum Gravity: A Progress Report", Rept.Prog.Phys.64:885,2001 [gr-qc/0108040]

[3] A. Ashtekar, "Quantum Mechanics of Geometry", gr-qc/9901023

[4] R. F. Streater, A. S. Wightman, "PCT, Spin and Statistics, and all that", Benjamin, New York, 1964

[5] R. Haag, "Local Quantum Physics", 2nd ed., Springer Verlag, Berlin, 1996

[6] S.A. Fulling, "Aspects of Quantum Field Theory in Curved Space-Time", Cambridge University Press, Cambridge, 1989
M. J. Radzikowski, "The Hadamard Condition and Kay's Conjecture in (Axiomatic) Quantum Field Theory on Curved Space-Time", PhD Thesis, Princeton University, October 1992

[7] R. Brunetti, K. Fredenhagen, "Microlocal Analysis and Interacting Quantum Field Theories", Commun.Math.Phys.208:623-661,2000, [math-ph/9903028]

[8] G. Scharf, "Finite Quantum Electrodynamics: The Causal Approach", Springer – Verlag, Berlin, 1995

[9] M.H. Goroff, A. Sagnotti, Nucl. Phys. **B266** (1986) 709

[10] P. Van Nieuwenhuizen, Phys. Rep. **68** (1981) 189

[11] N. Marcus, A. Sagnotti, Nucl. Phys. **B256** (1985) 77

[12] S. Deser, "Two Outcomes od Two Old (Super)problems, hep-th/9906178; "Infinities in Quantum Gravities", Annalen Phys.9:299-307,2000 [gr-qc/9911073]; "Non-Renormalizability of (Last Hope) $D = 11$ Supergravity with a Terse Survey of Divergences in Quantum Gravities, hep-th/9905017

[13] M.B. Green, J. Schwarz, E. Witten, "Superstring Theory", Cambridge University Press, Cambridge, 1987

[14] J. Polchinski, "String Theory", vol.1 : "An Introduction to the Bososnic String", vol. 2 : "Superstring Theory and Beyond", Cambridge University Press, Cambridge, 1998

[15] W. Lerche, "Recent Developments in String Theory, hep-th/9710246

[16] A. Bilal, "(M)atrix Theory : a pedagogical Introduction", hep-th/9710136
Washington Taylor IV, "Lectures on D-Branes, Gauge Theory and (M)atrices", hep-th/9801182

[17] M. Haack, B. Kors, D. Lüst, "Recent Developments in String Theory: From Perturbative Dualities to M Theory", hep-th/9904033
[18] I. Antoniadis, K. Benakli, M. Quiros, "Direct Collider Signatures of Large Extra Dimensions", Phys.Lett.B460:176-183,1999, [hep-ph/9905311]
[19] B. Zwiebach, "Closed String Field Theory: An Introduction", Les Houches Sum.Sch.1992:0647-678 (QC178:H6:1992), [hep-th/9305026]
[20] R. Helling, H. Nicolai, "Supermebranes and Matrix Theory", hep-th/9809103
[21] A. W. Peet, "TASI Lectures on Black Holes in String Theory", hep-th/0008241
[22] P.A.M Dirac, Phys. Rev. **73** (1948) 1092
P.A.M Dirac, Rev. Mod. Phys. **21** (1949) 392
[23] B. S. DeWitt, Phys. Rev. **160** (1967) 1113, Phys. Rev. **162** (1967) 1195, Phys. Rev. **162** (1967) 1239
J. A. Wheeler, "Geometrodynamics", Academic Press, New York, 1962
[24] P. A. M. Dirac, "Lectures on Quantum Mechanics", Belfer Graduate School of Science, Yeshiva University Press, New York, 1964
[25] A. Hanson, T. Regge, C. Teitelboim, "Constrained Hamiltonian Systems", Accdemia Nazionale dei Lincei, Roma, 1976
[26] M. Henneaux, C. Teitelboim, "Quantization of Gauge Systems" Princeton University Press, Princeton, 1992
[27] A. Ashtekar, M. Stillerman, "Geometric Quantization and Constrained Systems", Journ. Math. Phys. **27** (1986) 1319
[28] M. Gotay, "Constraints, Reduction and Quantization", J.Math.Phys.27:2051-2066,1986
[29] K. Kuchař, in : "Quantum Gravity II : A second Oxford Symposium", C.J. Isham, R. Penrose, D. W. Sciama (eds.), Clarendon Press, Oxford, 1981
[30] A. Ashtekar, Phys. Rev. Lett. **57** (1986) 2244, Phys. Rev. **D36** (1987) 1587
[31] J.B. Hartle, S. W. Hawking, Phys. Rev. **D28** (1983) 2960
[32] J. Glimm, A. Jaffe, "Quantum Physics", Springer-Verlag, New York, 1987
[33] M. Gell-Mann, J. Hartle, "Classical Equations for Quantum Systems" Phys.Rev.D47:3345-3382,1993, [gr-qc/9210010]
[34] R. Loll, "Discrete Approaches to Quantum Gravity in Four Dimensions", Living Rev.Rel.1:13,1998, [gr-qc/9805049]
[35] T. Regge, Nuovo Cimento **19** (1961) 558
R.M. Williams, P. Tuckey, Class. Quantum Grav. **9** (1992) 1409
R.M. Williams, Nucl. Phys. Procs. Suppl. **57** (1997) 73
[36] J. Ambjorn, M. Carfora, A. Marzuoli, "The Geometry of dynamical Triangulations", Springer-Verlag, Berlin, 1998
[37] J. Ambjorn, R. Loll, "Nonperturbative Lorentzian Quantum Gravity and Topology Change", Nucl.Phys.B536:407-434,1998, [hep-th/9805108]
J. Ambjorn, R. Loll, J.L. Nielsen, "Euclidean and Lorentzian Quantum Gravity: Lessons From Two Dimensions", Chaos Solitons Fractals 10:177-195,1999, [hep-th/9806241]
J. Ambjorn, K.N. Anagnostopoulos, R. Loll, "On the Phase Diagram of 2-D Lorentzian Quantum Gravity", Nucl.Phys.Proc.Suppl.83:733-735,2000, [hep-lat/9908054]
J. Ambjorn, K.N. Anagnostopoulos, R. Loll, "Crossing the C=1 Barrier in 2-D Lorentzian Quantum Gravity", Phys.Rev.D61:044010,2000, [hep-lat/9909129]
[38] J. Ambjorn, J. Correia, C. Kristjansen, R. Loll, "On the Relation Between Euclidean and Lorentzian 2-D Quantum Gravity", Phys.Lett.B475:24-32,2000, [hep-th/9912267]
J. Ambjorn, J. Jurkiewicz, R. Loll, "Lorentzian and Euclidean Quantum Gravity: Analytical

and Numerical Results", hep-th/0001124
J. Ambjorn, J. Jurkiewicz, R. Loll, "A Nonperturbative Lorentzian Path Integral for Gravity", Phys.Rev.Lett.85:924-927,2000, [hep-th/0002050]
J. Ambjorn, J. Jurkiewicz, R. Loll, "Nonperturbative 3-D Lorentzian Quantum Gravity", Phys.Rev.D64:044011,2001, [hep-th/0011276]
A. Dasgupta, R. Loll, "A Proper Time Cure For the Conformal Sickness in Quantum Gravity", Nucl.Phys.B606:357-379,2001, [hep-th/0103186]
[39] A. Connes, "Noncommutative Geometry", Academic Press, London, 1994
G. Landi, "An Introduction to Non-Commutative Spaces and Their Geometries", Springer Verlag, Berlin, 1997
[40] N. Seiberg, E. Witten : String Theory and Non-Comm. Geo "String Theory and Noncommutative Geometry", JHEP 9909:032,1999, [hep-th/9908142]
[41] C. Isham, "Topos Theory and Consistent Histories: The Internal Logic of the Set of all Consistent Sets", Int.J.Theor.Phys.36:785-814,1997, [gr-qc/9607069]
[42] R. Penrose, "Twistor Theory and the Einstein Vacuum", Class.Quant.Grav.16:A113-A130,1999
[43] D. Rideout, R. Sorkin, "Evidence for a Continuum Limit in Causal Set Dynamics", D.P. Rideout, R.D. Sorkin, Phys.Rev.D63:104011,2001, [gr-qc/0003117]
[44] By C.J. Isham, N. Linden, K. Savvidou, S. Schreckenberg, " Continuous Time and Consistent Histories", J.Math.Phys.39:1818-1834,1998, [quant-ph/9711031]
I. Kouletsis, Karel.V. Kuchar, "Diffeomorphisms as Symplectomorphisms in History Phase Space: Bosonic String Model", gr-qc/0108022
[45] A. Ashtekar with invited contributions, "New Perspectives in Canonical Gravity", Bibliopolis, Napoli, 1988
[46] R. Arnowitt, S. Deser, C.W.Misner, in : "Gravitation : An Introduction to current Research", L. Witten (ed.), Wiley, New York, 1962
[47] R. M. Wald, "General Relativity", The University of Chicago Press, Chicago, 1989
S. Hawking, Ellis, "The Large Scale Structure of Spacetime", Cambridge University Press, Cambridge, 1989
[48] R. Beig, O. Murchadha, Ann. Phys. **174** (1987) 463
[49] R. Geroch, "The Domain of Dependence", Journ. Math. Phys. **11** (1970) 437 - 509
[50] C.J. Isham, K. Kuchař, Ann. Phys. **164** (1985) 288
C.J. Isham, K. Kuchař, Ann. Phys. **164** (1985) 316
[51] N.M.J. Woodhouse, "Geometric Quantization", 2nd. edition, Clarendon Press, Oxford, 1991
[52] J.E. Marsden, P.R. Chernoff, "Properties of Infinite Dimensional Hamiltonian Systems", Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1974
[53] A. Ashtekar, "Non-Perturbative Canonical Gravity", Lectures notes prepared in collaboration with R.S. Tate, World Scientific, Singapore, 1991
A. Ashtekar, in : "Gravitation and Quantizations", B. Julia (ed.), Elsevier, Amsterdam, 1995
[54] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, T. Thiemann, "Quantization for diffeomorphism invariant theories of connections with local degrees of freedom", Journ. Math. Phys. **36** (1995) 6456-6493
[55] A. Higuchi Class. Quant. Grav. **8** (1991) 1983
A. Higuchi Class. Quant. Grav. **8** (1991) 2023
[56] A. Higuchi Class. Quant. Grav. **8** (1991) 1983; Class. Quant. Grav. **8** (1991) 2023
D. Marolf, "Path Integrals and Instantons in Quantum Gravity: Minisuperspace Models", Phys. Rev. **D53** (1996) 6979
D. Marolf, "Almost Ideal Clocks in Quantum Cosmology: A Brief Derivation of Time", Class.

Quantum Grav. **12** (1995) 2469-86
D. Marolf, "The spectral Analysis Inner Product for Quantum Gravity," in: Proceedings of the VIIth Marcel-Grossman Conference, R. Ruffini, M. Keiser (eds.), World Scientific, Singapore, 1995, [gr-qc/9409036]
D. Marolf, "Quantum Observable and Recollapsing Dynamics", Class. Quant. Grav. **12** (1995) 1199-1220, [gr-qc/9404053]
D. Marolf, Group Averaging and Refined Algebraic Quantization: Where Are We Now ?", in: proceedings of 9th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (MG 9), Rome, Italy, 2-9 Jul 2000, [gr-qc/0011112]
A. Gomberoff, D. Marolf , On Group Averaging for SO(N,1)" Int.J.Mod.Phys.D8:519-535,1999 [gr-qc/9902069]
D. Giulini, D. Marolf , "A Uniqueness Theorem for Constraint Quantization" Class.Quant.Grav.16:2489-2505,1999, [gr-qc/9902045]
D. Giulini, D. Marolf, "On the Generality of Refined Algebraic Quantization", Class.Quant.Grav.16:2479-2488,1999, [gr-qc/9812024]

[57] A. Ashtekar, A. Corichi, J.A. Zapata, "Quantum Theory of Geometry III : Non-Commutativity of Riemannian Structures", gr-qc/9806041

[58] Y. Yamasaki, "Measures on Infinite Dimensional Spaces", World Scientific, Singapore, 1985

[59] I. M. Gel'fand, N. Ya. Vilenkin, "Generalized Functions", vol. 4, Applications of Harmonic Analysis, Academic Press, New York, London, 1964

[60] D. Marolf, J. Mourão, T. Thiemann, Journ. Math. Phys. **38** (1997) 4370-40

[61] C. Rovelli, Class. Quantum Grav. **8** (1991) 297
C. Rovelli, Class. Quantum Grav. **8** (1991) 317

[62] C. Rovelli, Phys. Rev. **D43** (1990) 442
C. Rovelli, Phys. Rev. **D43** (1990) 2638

[63] A. Sen, Journ. Math. Phys. **22** (1981) 1781
A. Sen, Phys. Lett **B119** (1982) 89
A. Sen, Int. Journ. Theor. Phys. **21** (1982) 1

[64] R. Sexl, K. Urbantke, "Relativität, Gruppen, Teilchen", Springer-Verlag, Berlin, 1982
R. Penrose, W. Rindler, "Spinors and Spacetime' ' , vol. 1, vol. 2, Cambrige University Press, Cambridge, 1990

[65] R. Schoen, S.T. Yau, Comm. Math. Phys. **79** (1981) 231-60

[66] E. Witten, Comm. Math. Phys. **80** (1980) 381-402

[67] A. Ashtekar, in "Mathematics and General Relativity", American Mathematical Society, Providence, Rhode Island, 1987
A. Ashtekar, Contemporary Math **71** (1988) 39

[68] J. Samuel, Pramana Journ. Phys. **28** (1987) L429
T. Jacobson, L. Smolin, Phys. Lett. **B196** (1987) 39
T. Jacobson, L. Smolin, Class. Quantum Grav. **5** (1987) 583

[69] T. Jacobson, Class. Quantum Grav. **5** (1987) L143

[70] T. Jacobson, Class. Quantum Grav. **5** (1987) 923

[71] A. Ashtekar, J. D. Romano, R. S. Tate, Phys. Rev. **D40** (1989) 2572

[72] J. N. Goldberg, Phys. Rev. **D37** (1987) 2116

[73] M. Henneaux, J. E. Nelson, C. Schomblond, Phys. Rev. **D39** (1989) 434

[74] R. Capovilla, T. Jacobson, J. Dell, Class. Quantum Grav. **7** (1990) L1; Phys. Rev. Lett. **63** (1989) 2325; Class. Quantum Grav. **8** (1991) 59
R. Capovilla, T. Jacobson, J. Dell, L. Maison, Class. Quantum Grav. **8** 41

[75] P. Peldan, I. Bengtson, Phys. Lett. **B244** (1990) 261; In. Journ. Mod Phys. **A7** (1990) 261
P. Peldan, Class. Quantum Grav. **8** (1991) 1765

[76] F. Barbero, Phys. Rev. **D51** (1995) 5507
F. Barbero, Phys. Rev. **D51** (1995) 5498

[77] C. Rovelli, T. Thiemann, Phys. Rev. **D57** (1998) 1009-14
G. Immirzi, Nucl. Phys. Proc. Suppl. **57** (1997) 65
A. Corichi, K. Krasnov, Mod. Phys. Lett. **A13** (1998) 1339

[78] T. Thiemann, "Anomaly-free Formulation of non-perturbative, four-dimensional Lorentzian Quantum Gravity", Physics Letters **B380** (1996) 257-264, [gr-qc/9606088]

[79] B. C. Hall, Journ. Funct. Analysis **122** (1994) 103

[80] T. Thiemann, "Reality Conditions inducing Transforms for Quantum Gauge Field Theories and Quantum Gravity", T. Thiemann, Class. Quantum Gravity **13** (1996) 1383-1403, [gr-qc/9511057]

[81] T. Thiemann, "An Account of Transforms on $\mathcal{A}/\mathcal{G}$", Acta Cosmologica **21** (1995) 145-167, [gr-qc/9511049]

[82] A. Ashtekar, Phys. Rev. **D53** (Rapid Communications) (1996) R2865-69

[83] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, T. Thiemann, "Coherent State Transforms for Spaces of Connections", Journ. Funct. Analysis, **135** (1996) 519-551, [gr-qc/9412014]

[84] G.A. Mena Marugán, "Geometric Interpretation of Thiemann's Generalized Wick Transform", gr-qc/9705031
G.A. Mena Marugán, in : Proceedings of the Fifth Wigner Symposium, World Scientific, Singapore, to be published
G.A. Mena Marugán, "Thiemann Transform for Gravity with Matter Fields", gr-qc/9805010

[85] S. Holst, "Barbero's Hamiltonian Derived From a Generalized Hilbert-Palatini Action", Phys. Rev. **D53** (1996) 5966, gr-qc/9511026
N. Barros e Sá, "Hamiltonian Analysis of General Relativity with the Immirzi Parameter", gr-qc/0006013

[86] J. Samuel, "Canonical Gravity, Diffeomorphisms and Objective Histories", gr-qc/0005094; "Is Barbero's Hamiltonian Formulation a Gauge Theory of Lorentzian Gravity ?", Class. Quantum Grav. **17** (2000) L141, gr-qc/00050095

[87] T. Thiemann, "Generalized Boundary Conditions for General Relativity for the Asymptotically Flat Case in Terms of Ashtekar's Variables", Class. Quantum Grav. **12** (1995) 181-198, [gr-qc/9910008]

[88] M. Nakahara, "Geometry, Topology and Physics", IOP Publishing Ltd, Bristol, 1990

[89] J. W. Milnor, J. D. Staheff, "Characteristic Classes", Princeton University Press, Princeton, 1974

[90] R. Loll, "Imposing $\det(E) > 0$ in discrete Quantum Gravity", Phys. Lett. **B399** (1997) 227

[91] E. Binz, J. Sniatycki, H. Fischer, "Geometry of Classical Fields", North Holland, Amsterdam, 1988

[92] E. Hebey, "Sobolev Spaces on Riemannian manifolds", Lecture Notes in Mathematics **1635**, Springer Verlag, Berlin, 1996

[93] S. Lang, "Differential Manifolds", Addison Wesley, Reading, Massachusetts, 1972
Y. Choquet-Bruhat, C. DeWitt-Morette, "Analysis, Manifolds and Physics", part I, North Holland, Amsterdam, 1989

[94] P. R. Chernoff, J. E. Marsden, "Properties of Infinite Dimensional Hamiltonian Systems", Springer-Verlag, Berlin, 1974

[95] R. Gambini, A. Trias, Phys. Rev. **D22** (1980) 1380
C. Di Bartolo, F. Nori, R. Gambini, A. Trias, Lett. Nuov. Cim. **38** (1983) 497
R. Gambini, A. Trias, Nucl. Phys. **B278** (1986) 436

[96] R. Giles, Phys. Rev. **D8** (1981) 2160

[97] A. Ashtekar, J. Lewandowski, "Completeness of Wilson Loop Functionals on the Moduli Space of SL(2,C) and SU(1,1) Connections", Class. Quantum Grav. **10** (1993) L69, [gr-qc/9304044]

[98] T. Jacobson, L. Smolin, "Nonperturbative Quantum Geometries" Nucl. Phys. **B299** (1988) 295,
C. Rovelli, L. Smolin, "Loop Space Representation of Quantum General Relativity", Nucl. Phys. **B331**, 80 (1990)

[99] R. Gambini, J. Pullin, "Loops, Knots, Gauge Theories and Quantum Gravity", Cambridge University Press, Cambridge, 1996

[100] B. Brügmann, J. Pullin, Nucl. Phys. **B363** (1991) 221
B. Brügmann, J. Pullin, R. Gambini, Phys. Rev. Lett. **68** (1992) 431, Nucl. Phys. B **385** (1992) 1199

[101] L. H. Kauffman, "Knots and Physics", World Scientific, Singapore, 1991

[102] H. Kodama, Phys. Rev. **D42** (1990) 2548

[103] T. Thiemann, H.A. Kastrup, "Canonical Quantization of Spherically Symmetric Gravity in Ashtekar's Self-Dual Representation", Nucl. Phys. **B399** (1993) 211-258, [gr-qc/9310012]
T. Thiemann, "The Reduced Phase Space of Spherically Symmetric Einstein – Maxwell Theory Including a Cosmological Constant", Nucl. Phys. **B436** (1995) 681-720, [gr-qc/9910007]
T. Thiemann, "Reduced Models for Quantum Gravity", in: "Canonical Gravity: From Classical to Quantum", p. 289 - 318, Lecture Notes in Physics 434, J. Ehlers, H. Friedrich (Eds.), Springer-Verlag, Berlin 1994
A. Ashtekar, V. Husain, C. Rovelli, J. Samuel, L. Smolin, Class. Quantum Grav. **6** (1989) L185
V. Husain, L. Smolin, "Exactly Solvable Quantum Cosmologies From Two Killing Field Reductions of General Relativity", Nucl. Phys. **327** (1989) 205-238
H. Kodama, "Specialization of Ashtekar's Formalism to Bianchi Cosmology", Progr. Theor. Phys. **80**, (1988) 1024-1040
V. Husain, J. Pullin, "Quantum Theory of Spacetimes with One Killing Field", Mod. PHys. Lett. **A5** (1990) 733-741
A. Ashtekar, J. Pullin, "Bianchi Cosmologies: A New Description", Annals of the Israel Physical Society **9** (1990) 65-76
A. Ashtekar, V. Husain, "Symmetry Reduced Einstein Gravity and Generalized Sigma and Chiral Models", Int.J.Mod.Phys. **D7** 549-566, 1998, [gr-qc/9712053]

[104] A. Ashtekar, C.J. Isham, "Representations of the Holonomy Algebras of Gravity and Non-Abelean Gauge Theories", Class. Quantum Grav. **9** (1992) 1433, [hep-th/9202053]

[105] A. Rendall, "Comment on a Paper of Ashtekar and Isham", Class. Quantum Grav. **10** (1993) 605

[106] A. Ashtekar, J. Lewandowski, "Representation theory of analytic Holonomy $C^\star$ algebras", in "Knots and Quantum Gravity", J. Baez (ed.), Oxford University Press, Oxford 1994

[107] D. Marolf and J. M. Mourão, "On the Support of the Ashtekar-Lewandowski Measure", Commun. Math. Phys. **170** (1995) 583-606
A. Ashtekar, D. Marolf, J. Mourão, "Integration on the Space of Connections Modulo Gauge Transformations" in : Proceedings of the Lanczos International Centenary Conference, J. D. Brown, et al (eds.), SIAM, Philadelphia, 1994

[108] A. Ashtekar, J. Lewandowski, "Projective Techniques and Functional Integration for Gauge Theories", J. Math. Phys. **36**, 2170 (1995), [gr-qc/9411046]
J. Lewandowski, Int. J. Mod. Phys. **D3**, 207 (1994)

[109] A. Ashtekar, J. Lewandowski, "Differential Geometry on the Space of Connections via Graphs and Projective Limits", Journ. Geo. Physics **17** (1995) 191
J. Lewandowski, "Differential Geometry for the Space of Connections Modulo Gauge Transformations", in : Proceedings of the Lanczos International Centenary Conference, J. D. Brown, et al (eds.), SIAM, Philadelphia, 1994

[110] J. Baez, "Generalized Measures in Gauge Theory" Lett. Math. Phys. **31** (1994) 213-224, [hep-th/9310201]
J. Baez, "Diffeomorphism Invariant Generalized Measures on the Space of Connections Modulo Gauge Transformations", in: Proceedings of the conference "Quantum Topology", D. Yetter (ed.), World Scientific, Singapore, 1994, [hep-th/9305045]

[111] V. Husain, K. Kuchař, Phys. Rev. **D42** (1990) 12

[112] R. Penrose, in : "Quantum Gravity and Beyond", T. Bastin (ed.), Cambridge University Press, Cambridge, 1971

[113] C. Rovelli, L. Smolin, "Spin Networks and Quantum Gravity", Phys. Rev. **D53** (1995) 5743

[114] J. Baez, "Spin Networks in Non-Perturbative Quantum Gravity", in : "The Interface of Knots and Physics", L. Kauffman (ed.), American Mathematical Society, Providence, Rhode Island, 1996, [gr-qc/9504036]

[115] T. Thiemann, "The Inverse Loop Transform", Journ. Math. Phys. **39** (1998) 1236, [hep-th/9601105]

[116] J. Baez, S. Sawin, "Functional Integration on Spaces of Connections", q-alg/9507023
J. Baez, S. Sawin, "Diffeomorphism Invariant Spin-Network States", q-alg/9708005

[117] J. Lewandowski, T. Thiemann, "Diffeomorphism Invariant Quantum Field Theories of Connections in terms webs", Class. Quantum Grav. **16** (1999) 2299-2322, [gr-qc/9901015]

[118] J.-A. Zapata, "A Combinatoric Approach to Diffeomorphism Invariant Quantum Gauge Theories", Journ. Math. Phys. **38** (1997) 5663, [gr-qc/9703037]
J.-A. Zapata, "A Combinatoric Space from Loop Quantum Gravity", Gen. Rel. Grav. **30** (1998) 1229, [gr-qc/9703038]

[119] T. Thiemann, "Quantum Spin Dynamics (QSD) : III. Quantum Constraint Algebra and Physical Scalar Product in Quantum General Relativity", Class. Quantum Grav. **15** (1998) 1207-1247, [gr-qc/9705017]

[120] "Millenium Prize Problems", Clay Mathemematics Institute, http://www.claymath.org/prizeproblems/index.htm

[121] T. Balaban, J. Imbrie, A. Jaffe, "Exact Renormalization Group for Gauge Theories", in : Proceedings of the 1983 Cargèse Summer School
T. Balaban, A. Jaffe, "Constructive Gauge Theory", in : Proceedings of the 1986 Erichi Summer School
T. Balaban, "Constructive Gauge Theory II", G. Velo, A.S. Wightman (eds.), Plenum Press, New York, 1990
T. Balaban, Comm. Math. Phys. **89** (1983) 571, **95** (1984) 17, **96** (1984) 223, **98** (1985) 17, **99** (1985) 75, **99** (1985) 398, **102** (1985) 277, **109** (1987) 249, **116** (1988) 1, **119** (1988) 243, **122** (1989) 175, **122** (1989) 355
P. Federbush, "On the Quantum Yang-Mills Field Theory", Lecture Notes, Canadian Mathematical Society, Conference Proceedings, Vol. **9**, 1988
M. Göpfert, G. Mack, Comm. Math. Phys. **82** (1982) 545

E. Seiler, "Gauge Theory as a Problem in Constructive Quantum Field Theory and Statistical Mechanics", in : Lecture Notes in Physics, Springer-Verlag, Berlin, 1982
M. Asorey, P.K. Mitter, Commun. Math. Phys. **80** (1981) 43; M. Asorey, F. Falceto, Nucl Phys. **B327** (1989) 427
P. K. Mitter, P. Viallet, Comm. Math. Phys. **79**, 43 (1981)

[122] V. Rivasseau, "From perturbative to constructive Renormalization", Princeton University Press, Princeton, 1991

[123] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, T. Thiemann, in : "Geometry of Constrained Dynamical Systems", J. Charap (ed.), p. 60-86, Cambridge University Press, Cambridge, 1995
T. Thiemann, "An Axiomatic Approach to Constructive Quantum Gauge Field Theory", in : Symplectic Singularities and Geometry of Gauge Fields, Banach Center Publications, vol. 39, (1996) 389-403, [hep-th/9511122]

[124] K. Osterwalder, R. Schrader, Comm. Math. Phys. **31** (1973) 83, Comm. Math. Phys. **42** (1975) 281

[125] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, T. Thiemann, "$SU(N)$ Quatum Yang-Mills Theory in two dimensions : A complete solution", Journ. Math. Phys. **38** (1997) 5453-82, [hep-th/9605128]

[126] Y. M. Makeenko, A.A. Migdal, Phys. Lett. **B80** (1979) 255, Phys. Lett. **B97** (1980) 253, Nucl. Phys. **B153** (1979) 445, Yad. Phys. **32** (1980) 838, Yad. Phys. **33** (1980) 1639
Y. M. Makeenko, Yad. Phys. **33** (1980) 108
A. A. Migdal, Nucl. Phys. **B265** (1986) 594
L. Gross, C. King, A. Sengupta, Ann. Phys. **184** (1989) 65
B.K. Driver, Comm. Math. Phys. **123** (1989) 575
N. Bralic, Phys. Rev. **D22** (1980) 3090
S. Klimek, W. Kondracki, Commun. Math. Phys.**113** (1987) 389
E. Witten, Comm. Math. Phys. **141** (1991) 153, Journ. Geom. Phys. **9** (1991) 303
D. Gross, W. Taylor, Nucl. Phys. **B403** (1993) 395
M. Douglas, V. A. Kazakov, Phys. Lett. **B319** (1993) 219

[127] C. Fleischhack, "A New Type of Loop Independence and SU(N) Quantum Yang-Mills Theory in Two Dimensions", J.Math.Phys.40:2584-2610,1999, J.Math.Phys.41:76-102,2000; "On the Structure of Physical Measures in Gauge Theories"

[128] A. Ashtekar, D. Marolf, J. Mourão, T. Thiemann, "Constructing Hamiltonian Quantum Theories from Path Integrals in a Diffeomorphism Invariant Context", Class.Quant.Grav.17:4919-4940,2000, [quant-ph/9904094]

[129] M. Reed, B. Simon, "Methods of Modern Mathematical Physics: I. Functional Analysis", Academic Press, Boston, 1980

[130] J.M. Mourão, T. Thiemann, J.M. Velhinho, "Physical Properties of Quantum Field Theory Measures", J.Math.Phys.40:2337-2353,1999, [hep-th/9711139]

[131] J. Baez, "Spin Network States in Gauge Theory", Adv. Math. **117** (1996) 253-272

[132] J. Velhinho, "A Groupoid Approach to Spaces of Generalized Connections", hep-th/0011200

[133] C. Fleischhack , "Gauge Orbit Types for Generalized Connections", math-ph/0001006; "Hyphs and the Ashtekar – Lewandowski Measure", math-ph/0001007

[134] C. Fleischhack , "Stratification of the Generalized Gauge Orbit Space" math-ph/0001008; "On the Gribov Problem for Generalized Connections", math-ph/0007001

[135] J. M. Velhinho, "A Groupoid Approach to Spaces of Generalized Connections", hep-th/0011200
J.M. Velhinho, "Some Properties of Generalized Connections in Quantum Gravity", in: proceedings of the 9th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (MG 9), Rome, Italy, 2-9 Jul 2000, [hep-th/0101141]

[136] T. Thiemann, O. Winkler, "Gauge Field Theory Coherent States (GCS) : IV. Infinite Tensor Product and Thermodynamic Limit", accepted for publ. by Class. Quantum Grav., [hep-th/0005235]

[137] N. Biggs, "Algebraic Graph Theory", Cambridge University Press, 2nd ed., Cambridge, 1993
D. Stauffer, A. Aharony, "Introduction to Percolation Theory", Taylor and Francis, 2nd ed., London, 1994
D. M. Cvetovic, M. Doob, H. Sachs, "Spectra of Graphs", Academic Press, New York, 1979

[138] H. Sahlmann, T. Thiemann, O. Winkler, "Coherent States for Canonical Quantum General Relativity and the Infinite Tensor Product Extension", Nucl.Phys.B606:401-440,2001 [gr-qc/0102038]

[139] M. Nakahara, "Geometry, Topology and Physics", Institute of Physics Publishing, Bristol, 1998

[140] J. R. Munkres, "Toplogy: A First Course", Prentice Hall Inc., Englewood Cliffs (NJ), 1980

[141] J. L. Kelley, "General Topology", Springer-Verlag, Berlin, 1975

[142] R. V. Kadison, J. R. Ringrose, "Fundamentals of the Theory of Operator Algebras", vol. 1,2, Academic Press, London, 1983
O. Bratteli, D. W. Robinson, "Operator Algebras and Quantum Statistical Mechanics", vol. 1,2, Springer Verlag, Berlin, 1997
J. Dixmier, "Les Algebres d' Operateurs dans l' Espace Hilbertien (Algebres de von Neumann), Gauthiers-Villars, Paris, 1957

[143] W. Thirring, "Lehrbuch der Mathematischen Physik", vol. 3, Springer Verlag, Berlin, 1978
H. Baumgärtel, M. Wollenberg, "Causal Nets of Operator Algebras", Akademie Verlag, Berlin, 1992

[144] O. Landford III, "Selected Topoics in Functional Analysis", in: Proceedings of Les Houches Summer School "Statistical Mechanics and Quantum Field Theory", C. DeWitt, R. Stora (ed.), Gordon and Breach Science Publishers, London, 1971

[145] W. Rudin, "Real and Complex Analysis", McGraw-Hill, New York, 1987

[146] N. J. Vilenkin, "Special Functions and the Theory of Group Representations", American Mathematical Society, Providence, Rhode Island, 1968

[147] T. Thiemann, "Closed Formula for the Matrix Elements of the Volume Operator in Canonical Quantum Gravity", Journ. Math. Phys. **39** (1998) 3347-3371, [gr-qc/9606091]

[148] D. Giulini, "The Group of large Diffeomorphisms in Classical and Quantum Gravity", Published in Helv.Phys.Acta 69:333-334,1996
D. Giulini, "Determination and reduction of Large Diffeomorphisms", Nucl.Phys.Proc.Suppl.57:342-345,1997, [gr-qc/9702021]
D. Giulini, "The Group of Large Diffeomorphisms in General Relativity", Banach Center Publ.39:303-315,1997, [gr-qc/9510022]
D. Giulini, "Properties of Three Manifolds for Relativists", Int.J.Theor.Phys.33:913-930,1994, [gr-qc/9308008]

[149] H. Whitney, Ann. Math. **37** 648 – 680 (1936)

[150] C. Rovelli, L. Smolin, Nucl. Phys. **B442** (1995) 593, Erratum : Nucl. Phys. **B456** (1995) 734

[151] R. Loll, "Spectrum of the Volume Operator in Quantum Gravity", Nucl.Phys.B460:143-154,1996 [gr-qc/9511030]

[152] R. Loll, Class. Quantum Grav. **14** (1997) 1725
R. Loll, Nucl. Phys. **B500** (1997) 405
R. Loll, Phys. Rev. Lett. **75** (1995) 3048
[153] A. Ashtekar, J. Lewandowski, "Quantum Theory of Geometry I: Area Operators", Class. Quantum Grav. **14** (1997) A55-81 A. Ashtekar, J. Lewandowski, "Quantum Theory of Geometry II : Volume Operators", Adv. Theo. Math. Phys. **1** (1997) 388-429
J. Lewandowski, Class. Quantum Grav. **14** (1997) 71-76
[154] R. De Pietri, C. Rovelli, "Geometry Eigenvalues and Scalar Product from Recoupling Theory in Loop Quantum Gravity", Phys. Rev. **D54** (1996) 2664, [gr-qc/9602023]
[155] R. De Pietri, Class. Quantum Grav. **14** (1997) 53
[156] T. Thiemann, "A Length Operator for Canonical Quantum Gravity", Journ. Math. Phys. **39** (1998) 3372-3392, [gr-qc/9606092]
[157] T. Thiemann, "Quantum Spin Dynamics (QSD) : VI. Quantum Poincaré Algebra and a Quantum Positivity of Energy Theorem for Canonical Quantum Gravity", Class. Quantum Grav. **15** (1998) 1463-1485, [gr-qc/9705020]
[158] S. Major, "Modeling Space with an Atom of Geometry", gr-qc/0109056,
S. Major, "Operators for Quantized Directions", Class.Quant.Grav.16:3859-3877,1999, [gr-qc/9905019]
[159] Y. Ma, Y. Ling, "The Q-Hat Operator for Canonical Quantum Gravity", Phys.Rev.D62:104021,2000, [gr-qc/0005117]
[160] A. R. Edmonds, "Angular Momentum in Quantum Mechanics", Princeton University Press, Princeton, 1974
[161] A. Alekseev, A.P. Polychronakos, M. Smedback, "On Area and Entropy of a Black Hole", [hep-th/0004036]
[162] S. Fairhurst, "Table of Lowest Hundred Eigenvalues for the Area Operator", unpublished
[163] T. Thiemann, "A Diffeomorphism Invariant Volume Operator", in preparation
[164] T. Thiemann, "Quantum Spin Dynamics (QSD)", Class. Quantum Grav. **15** (1998) 839-73, [gr-qc/9606089]
[165] T. Thiemann, "Quantum Spin Dynamics (QSD) : II. The Kernel of the Wheeler-DeWitt Constraint Operator", Class. Quantum Grav. **15** (1998) 875-905, [gr-qc/9606090]
[166] C. Di Bartolo, R. Gambini, J. Griego, J. Pullin, "Canonical Quantum Gravity in the Vasiliev Invariants Arena: I. Kinematical Structure", Class. Quantum Grav. **17** (2000) 3211-3237
[167] C. Di Bartolo, R. Gambini, J. Griego, J. Pullin, "Canonical Quantum Gravity in the Vasiliev Invariants Arena: II. Constraints, Habitats and Consistency of the Constraint Algebra", Class. Quantum Grav. **17** (2000) 3239-3264
[168] M. Bobienski, J. Lewandowski, M. Mroczek, "A Two Surface Quantization of Lorentzian Gravity", gr-qc/0101069
[169] M. Blencowe, Nucl. Phys. **B341** (1990) 213
[170] B. Brügmann, J. Pullin, Nucl. Phys. **B390** (1993) 399
[171] R. Gambini, Phys. Lett. **B255** (1991) 180
R. Gambini, A. Garat, J. Pullin, Int. Journ. Mod. Phys. **D4** (19995) 589
[172] C. Rovelli, L. Smolin, Phys. Rev. Lett. **72** (1994) 446
[173] C. Rovelli, Nucl. Phys. **B405** (1993) 797
L. Smolin, Phys. Review **D49** (1994) 4028
[174] R. Loll, Acta Cosmologica **21** (1995) 132
R. Loll, Phys. Rev. **D54** (1996) 5381

[175] J. Baez, in : "Matters of Gravity 1996", webpage html://www.phys.psu.edu/PULLIN/, gr-qc/9609008
[176] M. Creutz, "Quarks, Gluons and Lattices", Cambridge University Press, Cambridge, 1983
[177] P. Hajíček, K. Kuchař, Phys. Rev. **D41** (1990) 1091, Journ. Math. Phys. **31** (1990) 1723
[178] M. Gaul, C. Rovelli, "A Generalized Hamiltonian Contraint Operator in Loop Quantum Gravity and its Simplest Euclidean Matrix Elements", Class.Quant.Grav.18:1593-1624,2001, [gr-qc/0011106]
[179] D. Marolf, J. Lewandowski, "Loop Constraints : A Habitat and their Algebra", Int.J.Mod.Phys.D7:299-330,1998, [gr-qc/9710016]
[180] R. Gambini, J. Lewandowski, D. Marolf, J. Pullin, "On the Consistency of the Constraint Algebra in Spin Network Gravity", Int.J.Mod.Phys.D7:97-109,1998, [gr-qc/9710018]
[181] T. Thiemann, "Quantum Spin Dynamics (QSD): VIII. The Classical Limit", in preparation
[182] L. Smolin, "The classical Limit and the Form of the Hamiltonian Constraint in Non-Perturbative Quantum General Relativity", gr-qc/9609034
[183] R. Borissov, R. De Pietri, C. Rovelli, "Matrix Elements of Thiemann's Hamiltonian Constraint in Loop Quantum Gravity", Class. Quantum Grav. **14** (1997) 2793-2823
[184] T. Thiemann, "Quantum Spin Dynamics (QSD) : IV. 2+1 Euclidean Quantum Gravity as a model to test 3+1 Lorentzian Quantum Gravity", Class. Quantum Grav. **15** (1998) 1249-1280, [gr-qc/9705018]
[185] M. Bojowald, H. Kastrup, "Quantum Symmetry Reduction for Diffeomorphism Invariant Theories of Connections", Class. Quantum Grav. **17** (2000) 3009 [hep-th/9907042]; "The Area Operator in the Spherically Symmetric Sector of Loop Quantum Gravity", Class. Quantum Grav. **17** (2000) 3044 [hep-th/9907043]
M. Bojowald, "Abelean BF Theory and Spherically Symmetric Electromagnetism", J. Math. Phys. **41** (2000) 4313 [hep-th/9908170]
M. Bojowald, "Loop Quantum Cosmology. I. Kinematics", Class. Quantum Grav. **17** (2000) 1489 [gr-qc/9910103]
M. Bojowald, "Loop Quantum Cosmology. II. Volume Operators", Class. Quantum Grav. **17** (2000) 1509 [gr-qc/9910104]
M. Bojowald, "Loop Quantum Cosmology. III. Wheeler-DeWitt Operators", Class.Quant.Grav.18:1055-1070,2001 [gr-qc/0008052]
M. Bojowald, "Loop Quantum Cosmology. IV. Discrete Time Evolution" Class.Quant.Grav.18:1071-1088,2001 [gr-qc/0008053]
M. Bojowald, "Absence of Singularity in Loop Quantum Cosmology", Phys.Rev.Lett.86:5227-5230,2001, [gr-qc/0102069]
M. Bojowald, "Dynamical Initial Conditions in Quantum Cosmology", Phys.Rev.Lett.87:121301,2001, [gr-qc/0104072]
M. Bojowald, "The Inverse Scale Factor in Isotropic Quantum Geometry", Phys.Rev.D64:084018,2001, [gr-qc/0105067]
M. Bojowald, "The Semiclassical Limit of Loop Quantum Cosmology", Class.Quant.Grav.18:L109-L116,2001, [gr-qc/0105113]
[186] S. Carlip, "Lectures in (2+1)-dimensional gravity", gr-qc/9503024
S. Carlip, "The Statistical Mechanics of the three-dimensional Euclidean Black Hole", Phys. Rev. **D55** (1997) 878-882
[187] E. Witten, Nucl. Phys. **B311** (1988) 46

[188] A. Mikovic, N. Manojlovic, Nucl. Phys. **B385** (1992) 571
A. Mikovic, N. Manojlovic, "Remarks on the Reduced Phase Space of (2+1) Gravity on a Torus in the Ashtekar Formulation", gr-qc/9712011
[189] F. Barbero, M. Varadarajan, Nucl. Phys. **B415** (1994) 515
[190] D. Marolf, Class. Quantum Grav. **10** (1993) 2625
D. Marolf, Canadian Gen. Rel. **14** (1993) 256
J. Louko, D. Marolf, Class. Quantum Grav. **11** (1994) 311
[191] A. Ashtekar, R. Loll, Class. Quantum Grav. **11** (1994) 2417
R. Loll, Class. Quantum Grav. **12** (1995) 1655
[192] T. Jacobson, "(1+1) Sector of (3+1) Gravity", Class.Quant.Grav.13:L111-L116,1996, Erratum-ibid.13:3269,1996, [gr-qc/9604003]
J. Lewandowski, Jacek Wisniewski, "Degenerate Sectors of Ashtekar Gravity", Class.Quant.Grav.16:3057-3069,1999, [gr-qc/9902037]
[193] T. Thiemann, "Quantum Spin Dynamics (QSD) : V. Quantum Gravity as the Natural Regulator of the Hamiltonian Constraint of Matter Quantum Field Theories", Class. Quantum Grav. **15** (1998) 1281-1314, [gr-qc/9705019]
[194] T. Thiemann, "Kinematical Hilbert Spaces for Fermionic and Higgs Quantum Field Theories", Class. Quantum Grav. **15** (1998) 1487-1512, [gr-qc/9705021]
[195] C. Rovelli, H. Morales-Tecótl, Phys. Rev. Lett **72** (1995) 3642-3645
C. Rovelli, H. Morales-Tecótl, Nucl. Phys. **B331** (1995) 325-361
[196] J. Baez, K. Krasnov, "Quantization of Diffeomorphism Invariant Theories with Fermions", Journ. Math. Phys **39** (1998) 1251
[197] Y. Choquet-Bruhat, C. DeWitt-Morette, "Analysis, Manifolds and Physics", part II, North Holland, Amsterdam, 1989
[198] B. DeWitt, "Supermanifolds", Cambridge University Press, Cambridge, 1992
[199] I. L. Buchbinder, S.L. Lyahovich, Class. Quantum Grav. **4** (1987) 1487
[200] D.M. Gitman, I. V. Tyutin, "Quantization of Fields with Constraints", Springer-Verlag, Berlin, 1990
[201] A. Ashtekar, C. Rovelli, L. Smolin, "Weaving a Classical Geometry with Quantum Threads", Phys.Rev.Lett.69:237-240,1992, [hep-th/9203079]
J. Zegwaard, "Weaving of Curved Geometries", Phys.Lett.B300:217-222,1993, [hep-th/9210033]
J. Zegwaard, "Gravitons in Loop Quantum Gravity", Nucl.Phys.B378:288-308,1992
J. Iwasaki, C. Rovelli, "Gravitons as Embroidery of the Weave", Int.J.Mod.Phys.D1:533-557,1993
J. Iwasaki, C. Rovelli, "Gravitons from Loops: Nonperturbative Loop Space Quantum Gravity Contains The Graviton Physics Approximation", Class.Quant.Grav.11:1653-1676,1994
Y. Ma, Y. Ling, "The Classical Geometry from a Physical State in Canonical Qunatum Gravity", Phys.Rev.D62:064030,2000, [gr-qc/0004070]
[202] M. Arnsdorf, Sameer Gupta, "Loop Quantum Gravity on Noncompact Spaces", Nucl.Phys.B577:529-546,2000, [gr-qc/9909053]
M. Arnsdorf, "Approximating Connections in Loop Quantum Gravity", gr-qc/9910084
[203] M. Varadarajan, J. A. Zapata, "A proposal for Analyzing the Classical Limit of Kinematic Loop Gravity", Class.Quant.Grav.17:4085-4110,2000, [gr-qc/0001040]
[204] L. Bombelli, "Statistical Lorentzian Geometry and the Closeness of Lorentzian Manifolds, J.Math.Phys.41:6944-6958,2000, [gr-qc/0002053]

[205] C. Itzykson, J.-M.Drouffe, "Statistical Field Theory", vol. 2, Cambridge University Press, Cambridge, 1989

[206] M. Bojowald, private communication

[207] T. Thiemann, "Quantum Spin Dynamics (QSD): VII. Symplectic Structures and Continuum Lattice Formulations of Gauge Field Theories", T. Thiemann, Class.Quant.Grav.18:3293-3338,2001, [hep-th/0005232]
T. Thiemann, "Gauge Field Theory Coherent States (GCS): I. General Properties", Class.Quant.Grav.18:2025-2064,2001, [hep-th/0005233]

[208] T. Thiemann, O. Winkler, "Gauge Field Theory Coherent States (GCS): II. Peakedness Properties", Class.Quant.Grav.18:2561-2636,2001, [hep-th/0005237]
T. Thiemann, O. Winkler, "Gauge Field Theory Coherent States (GCS): III. Ehrenfest Theorems", accepted for publ. by Class. Quantum Grav., [hep-th/0005234]

[209] T. Thiemann, O. Winkler, "Gauge Field Theory Coherent States (GCS): IV. Infinite Tensor Product and Thermodynamical Limit", accepted for publ. by Class. Quantum Grav., hep-th/0005235

[210] H. Sahlmann, T. Thiemann, O. Winkler, " Coherent States for Canonical Quantum General Relativity and the Infinite Tensor product Extension", Nucl.Phys.B606:401-440,2001, [gr-qc/0102038]

[211] H. Sahlmann, T. Thiemann, "Quantum Field Theory on Curved Spacetimes from Canonical Quantum General Relativity (QFTQGR) : I. Scalar Fields, II. Photons and the γ Ray Burst Effect, III. Gravitons, IV. Leptons", in preparation

[212] M. Varadarajan, "Fock representations from U(1) Holonomy Algebras", Phys.Rev.D61:104001,2000, [gr-qc/0001050]
M. Varadarajan, "Photons from Quantized Electric Flux Representations", gr-qc/0104051
M. Varadarajan, "Photons from Quantized Electric Flux Representations", in preparation

[213] A. Ashtekar, C. Isham, "Inequivalent Observable Algebras: Another Ambiguity in Field Quantization", Phys.Lett.B274:393-398,1992
A. Ashtekar, Carlo Rovelli, L. Smolin, "Gravitons and Loops", Phys.Rev.D44:1740-1755,1991 [hep-th/9202054]

[214] J. Velhinho, "Invariance Properties of Induced Fock Measures for U(1) Holonomies", math-ph/0107002

[215] A. Ashtekar, J. Lewandowski, "Relation Between Polymer and Fock Excitations", Class.Quant.Grav.18:L117-L128,2001, [gr-qc/0107043]

[216] J. Klauder, B.-S. Skagerstam, "Coherent States", World Scientific, Singapore, 1985
A. Perelomov, "Generalized Coherent States and Their Applications", Springer-Verlag, Berlin, 1986

[217] F. Bayen, M. Flato, C. Fronsdal, A. Liechnerowicz, D. Sternheimer: "Deformation Theory and Quantization", Annals of Physics **111** (1978) 61-110; Annals of Physics **111** (1978) 111-151;

[218] H. Sahlmann, T. Thiemann, "Coherent Distributions for Canonical Quantum General Relativity", in preparation

[219] J. von Neumann, Comp. Math. **6** (1938) 1-77

[220] W. Thirring, "Lehrbuch der Mathematischen Physik", vol. 4, Springer-Verlag, Wien, 1994

[221] G. Amelino-Camelia, "Are we at Dawn with Quantum Gravity Phenomenology", Lectures given at 35th Winter School of Theoretical Physics: From Cosmology to Quantum Gravity, Polanica, Poland, 2-12 Feb 1999, Lect.Notes Phys.541:1-49,2000, [gr-qc/9910089]
G. Amelino-Camelia, "Gravity Wave Interferometers as Probes of a Low Energy Effective Quantum Gravity", Phys.Rev.D62:024015,2000 [gr-qc/9903080]

G. Amelino-Camelia, "An Interferometric Gravitational Wave Detector as a Quantum Gravity Apparatus", Nature 398:216-218,1999, [gr-qc/9808029]
G. Amelino-Camelia, John R. Ellis, N.E. Mavromatos, D.V. Nanopoulos, Subir Sarkar, "Potential Sensitivity of Gamma Ray Burster Observations to Wave Dispersion in Vacuo", Nature 393:763-765,1998 [astro-ph/9712103]

[222] S. Carlip, "Quantum Gravity: A Progress Report", Rept.Prog.Phys.64:885,2001 [gr-qc/0108040]

[223] R. Gambini, J. Pullin, "Nonstandard Optics from Quantum Spacetime" Phys.Rev.D59:124021,1999, [gr-qc/9809038]
R. Gambini, J. Pullin, "Quantum Gravity Experimental Physics ?", Gen.Rel.Grav.31:1631-1637,1999

[224] J. Alfaro, H. A. Morales-Tecotl, L. F. Urrutia, "Quantum Gravity Corrections to Neutrino Propagation", Phys.Rev.Lett.84:2318-2321,2000, [gr-qc/9909079]
J. Alfaro, H. A. Morales-Tecotl, L. F. Urrutia, "Loop Quantum Gravity and Light Propagation", hep-th/0108061

[225] S. D. Biller et al, Phys. Rev. Lett. **83** (1999) 2108

[226] J. C. Baez , "An Introduction to Spin Foam Models of Quantum Gravity and BF Theory", Lect.Notes Phys.543:25-94,2000, [gr-qc/9905087]
J. C. Baez, "Spin Foam Models", Class.Quant.Grav.15:1827-1858,1998, [gr-qc/9709052]

[227] J. W. Barrett, "State Sum Models for Quantum Gravity", gr-qc/0010050
J. W. Barrett, "Quantum Gravity as Topological Quantum Field Theory", J.Math.Phys.36:6161-6179,1995, [gr-qc/9506070]

[228] A. Ashtekar, C. Beetle, O. Dreyer, S. Fairhurst, B. Krishnan, J. Lewandowski, J. Wisniewski "Isolated Horizons and Their Applications", Phys.Rev.Lett.85:3564-3567,2000 [gr-qc/0006006]
A. Ashtekar, "Classical and Quantum Physics of Isolated Horizons", Lect.Notes Phys.541:50-70,2000
A. Ashtekar, "Interface of general Relativity, Quantum Physics and Statistical Mecahnics: Some recent Developments", Annalen Phys.9:178-198,2000, [gr-qc/9910101]
A. Ashtekar, C. Beetle, S. Fairhurst, " Isolated Horizons: A Generalization of Black Hole Mechanics", Class.Quant.Grav.16:L1-L7,1999, [gr-qc/9812065]
A. Ashtekar, K. Krasnov, "Quantum Geometry and Black Holes", gr-qc/9804039

[229] A. Ashtekar, A. Corichi, K. Krasnov, "Isolated Horizons: The Classical Phase Space", Adv.Theor.Math.Phys.3:419-478,2000, [gr-qc/9905089]

[230] A. Ashtekar, J. C. Baez, K. Krasnov, "Quantum Geometry of Isolated Horizons and Black Hole Entropy", Adv.Theor.Math.Phys.4:1-94,2001, [gr-qc/0005126]

[231] G. Ponzano, T. Regge, "Semiclassical Limit of Racah Coefficients", in: Spectroscopy and Group Theoretical Methods in Physics", ed. F. Bloch, North Holland, New York, 1968
V. Turarev, O. Viro, "State Sum Invariants of 3-Manifolds and Quantum 6j Symbols", Topology **31** (1992) 865-902
H. Ooguri, "Topological Lattice Models in Four Dimensions", Mod. Phys. Lett. **A7** (1992) 2799-2810
L. Crane, D. Yetter, "A Categorical Construction of 4D TQFTs", in: "Quantum Topology", p. 120-130, eds. L. Kauffman, R. Baadhio, World Scientific, Singapore, 1993
L. Crane, L. Kauffman, D. Yetter, "State-Sum Invariants of 4-Manifolds", J. Knot Theory & Ramifications **6** (1997) 177-234

[232] M. F. Atiyah, “Topological Quantum Field Theories”, Publ. Math. IHES **68** (1989) 175-186
M. F. Atiyah, “The Geometry of Physics and Knots”, Cambridge University Press, Cambridge 1990
C. Kassel, “Quantum Groups”, Springer Verlag, Berlin, 1995
L. Kauffman, “Knots and Physics”, World Scientific Press, Singapore, 1993
L. Kauffman, S. Lins, “Temperley-Lieb recoupling Theory and Invariants of 3-Manifolds”, Princeton University Press, Princeton, 1994
V. Tuarev, “Quantum Invariants of Knots and 3-Manifolds”, de Gruyter, New York, 1994
E. Witten, “Quantum Field Theory and the Jones Polynomial”, Comm. Math. Phys. **121** (1989) 351-399
N. Reshetikhin, “Invariants of 3-manifolds via link polynomials and quantum groups”, Invent. Math. **103** 547-597 (1991)
D. Birmingham, M. Blau, M. Rakowski, G. Thompson, “Topological Field Theories”, Phys. Rep. **209** (1991) 129-340
R. Friedman, J. Morgan, “Gauge Theory and the Topology of Four – Manifolds”, AMS, Providence, 1998

[233] M. Reisenberger, C. Rovelli, “Sum over Surfaces Form of Loop Quantum Gravity”, Phys. Rev. **D56** (1997) 3490-3508

[234] M. Reisenberger, “World Sheet Formulations of Gauge Theories and Gravity”, [gr-qc/9412035]
M. P. Reisenberger, “A Lefthanded Simplicial Action for Euclidean General Relativity”, Class.Quant.Grav.14:1753-1770,1997, [gr-qc/9609002]

[235] L. Freidel, K. Krasnov, R. Puzio, “BF Description of Higher Dimensional Gravity Theories”, Adv.Theor.Math.Phys.3:1289-1324,1999, [hep-th/9901069]

[236] L. Freidel, K. Krasnov, “Spin Foam Models and the Classical Action Principle”, Adv.Theor.Math.Phys.2:1183-1247,1999, [hep-th/9807092]

[237] A. Barbieri, “Space of Vertices of Relativistic Spin Networks”, gr-qc/9709076
A. Barbieri, “Quantum Tetrahedra and Simplicial Spin Networks”, Nucl.Phys.B518:714-728,1998, [gr-qc/9707010]

[238] J. C. Baez, J. W. Barrett, “The Quantum Tetrahedron in Three Dimensions and Four Dimensions”, Adv.Theor.Math.Phys.3:815-850,1999, [gr-qc/9903060]

[239] J. W. Barrett, L. Crane, “Relativistic Spin Networks and Quantum Gravity”, J.Math.Phys.39:3296-3302,1998, [gr-qc/9709028]

[240] M. P. Reisenberger, “On Relativistic Spin Network Vertices”, J.Math.Phys.40:2046-2054,1999, [gr-qc/9809067]

[241] D. Yetter, “Generalized Barrett-Crane Vertices and Invariants of Embedded Graphs”, qa/9801131
J. W. Barrett, “The Classical Evaluation of Relativistic Spin Networks”, Adv.Theor.Math.Phys.2:593-600,1998, [math.qa/9803063]
J. W. Barrett, R. M. Williams, “The Asymptotics of an Amplitude for the Four Simplex”, Adv.Theor.Math.Phys.3:209-215,1999, [gr-qc/9809032]

[242] J. Ambjorn, B. Durhuus, T. Jonnson, “Quantum Geometry: A Statistical Field Theory Approach”, Cambridge University Press, Cambridge, 1997

[243] D. V. Boulatov, Mod. Phys. Lett. **A7** (1992) 1629
H. Ooguri, Mod. Phys. Lett. **A7** (1992) 2799

[244] R. De Pietri, L. Freidel, K. Krasnov, C. Rovelli, “Barrett-Crane Model from a Boulatov-Ooguri Field Theory Over a Homogeneous Space”, Nucl.Phys.B574:785-806,2000, [hep-th/9907154]

[245] A. Perez, C. Rovelli, "A Spin Foam Model Without Bubble Divergences", Nucl.Phys.B599:255-282,2001, [gr-qc/0006107]
[246] A. Perez, "Finiteness of a Spin Foam MOdel for Euclidean Quantum General Relativity", Nucl.Phys.B599:427-434,2001, [gr-qc/0011058]
[247] M. Reisenberger, C. Rovelli, "Spin Foams as Feynman Diagrams", gr-qc/0002083
M. P. Reisenberger, C. Rovelli, "Space Time as a Feynman Diagram: The Connection Formulation", Class.Quant.Grav.18:121-140,2001, [gr-qc/0002095]
[248] J. W. Barrett, L. Crane, "A Lorentzian Signature Model for Quantum General Relativity", Class.Quant.Grav.17:3101-3118,2000, [gr-qc/9904025]
[249] I. M. Gel'fand, M. A. Naimark, "Unitary Representations of the Proper Lorentz Group", Izv. Akad. Nauk. SSSR **11** (1947) 411
[250] J. C. Baez, J. W. Barrett, "Integrability of Relativistic Spin Networks", gr-qc/0101107
[251] A. Perez, C. Rovelli, "Spin Foam Model for Lorentzian General Relativity", Phys.Rev.D63:041501,2001, [gr-qc/0009021]
[252] L. Crane, A. Perez, C. Rovelli, "A Finiteness Proof for the Lorentzian State Sum Spin Foam Model for Quantum General Relativity", gr-qc/0104057
[253] A. Perez, C. Rovelli, "Observables in Quantum Gravity", gr-qc/0104034
[254] F. Markopoulou, L. Smolin, "Causal Evolution of Spin Networks", Nucl.Phys.B508:409-430,1997, [gr-qc/9702025]
F. Markopoulou, "Dual Formulation of Spin Network Evolution", gr-qc/9704013
F. Markopoulou, L. Smolin, "Quantum Geometry with Intrinsic Local Causality", Phys.Rev.D58:084032,1998, [gr-qc/9712067]
F. Markopoulou, "The Internal Description of a Causal Set: What the Universe Like from Inside", Commun.Math.Phys.211:559-583,2000, [gr-qc/9811053]
F. Markopoulou, "Quantum Causal Histories", Class.Quant.Grav.17:2059-2072,2000, [hep-th/9904009]
F. Markopoulou, "An Insider's Guide to Quantum Causal Histories", Nucl.Phys.Proc.Suppl.88:308-313,2000, [hep-th/9912137]
[255] F. Markopoulou, "An Algebraic Approach to Coarse Graining", hep-th/0006199
[256] A. Connes, D. Kreimer, "Renormalization in Quantum Field Theory and the Riemann – Hilbert Problem", JHEP 9909:024,1999 [hep-th/9909126]
A. Connes, D. Kreimer, "Renormalization in Quantum Field Theory and the Riemann – Hilbert Problem. 1. The Hopf Algebra Structure of Graphs and the Main Theorem", Commun.Math.Phys.210:249-273,2000, [hep-th/9912092]
A. Connes, D. Kreimer, "Renormalization in Quantum Field Theory and the Riemann – Hilbert Problem. 2. The Beta Function, Diffeomorphisms and the renormalization Group", Commun.Math.Phys.216:215-241,2001, [hep-th/0003188]
[257] J. D. Bekenstein, "Black Holes and Entropy", Phys.Rev. **D7** (1973) 2333-2346
J. D. Bekenstein, "Generalized Second Law for Thermodynamics in Balck Hole Physics", Phys. Rev. **D9** (1974) 3292-3300
[258] K. Krasnov, "On Statistical Mechanics of Gravitational Systems", Gen.Rel.Grav.30:53-68,1998, [gr-qc/9605047]
[259] C. Rovelli, "Black Hole Entropy from Loop Quantum Gravity" Phys.Rev.Lett.77:3288-3291,1996, [gr-qc/9603063]

[260] S. Axelrod, S. D. Pietra, E. Witten, "Geometric Quantization of Chern-Simons Gauge Theory", J. Diff. Geo. **33** (1991) 787-902
[261] D. Mumford, "Tata Lectures on Theta I", Birkäuser, Boston, 1983
[262] A. Ashtekar, A. Corichi, "Laws Governing Isolated Horizons: Inclusion of Dilaton Couplings", Class.Quant.Grav.17:1317-1332,2000, [gr-qc/9910068]
A. Ashtekar, A. Corichi, D. Sudarsky, "Hairy Black Holes, Horizon Mass and Solitons", Class.Quant.Grav.18:919-940,2001, [gr-qc/0011081]
S. Fairhurst, B. Krishnan, "Distorted Balck Holes with Charge", gr-qc/0010088
[263] A. Ashtekar, C. Beetle, S. Fairhurst, "Mechanics of Isolated Horizons", Class.Quant.Grav.17:253-298,2000, [gr-qc/9907068]
A. Ashtekar, S. Fairhurst, B. Krishnan, "Isolated Horizons: Hamiltonian Evolution and the First Law", Phys.Rev.D62:104025,2000, [gr-qc/0005083]
[264] A. Ashtekar, C. Beetle, J. Lewandowski, "Mechanics of Rotating Isolated Horizons", Phys.Rev.D64:044016,2001, [gr-qc/0103026]
[265] M. Barreira, M. Carfora, C. Rovelli, "Physics with Nonperturbative Quantum Gravity: Radiation from a Black Hole", Gen.Rel.Grav.28:1293-1299,1996, [gr-qc/9603064]
[266] J. Bekenstein, V. Mukhanov, "Spectroscopy of the Quantum Black Hole", Phys.Lett.B360:7-12,1995, [gr-qc/9505012]
[267] K. Krasnov, "Quantum Geometry and Thermal Radiation from Black Holes", Class.Quant.Grav.16:563-578,1999, [gr-qc/9710006]
[268] S. Carlip, "Entropy from Conformal Field Theory at Killing Horizons", Class.Quant.Grav.16:3327-3348,1999, [gr-qc/9906126] S. Carlip, "Black Hole Entropy from Horizon Conformal Field Theory", Nucl.Phys.Proc.Suppl.88:10-16,2000, [gr-qc/9912118]
[269] O. Dreyer, A. Ghosh, J. Wisniewski, "Black Hole Entropy Calculations Based on Symmetries", Class.Quant.Grav.18:1929-1938,2001, [hep-th/0101117]
[270] L. Smolin, "Strings as Perturbations of Evolving Spin Networks", Nucl.Phys.Proc.Suppl.88:103-113,2000, [hep-th/9801022]
L. Smolin, "A Holographic Formulation of Quantum General Relativity", Phys.Rev.D61:084007,2000, [hep-th/9808191]
L. Smolin, "Towards a Background Independent Approach to M Theory", hep-th/9808192
S. Kauffman, L. Smolin, "Combinatorical Dynamics in Quantum Gravity", Published in Lect.Notes Phys.541:101-129,2000, [hep-th/9809161]
L. Smolin, "A Candidate for a Background Independent Formulation of M Theory", Phys.Rev.D62:086001,2000, [hep-th/9903166]
Y. Ling, L. Smolin, "Supersymmetric Spin Networks and Quantum Supergravity", Phys.Rev.D61:044008,2000, [hep-th/9904016]
F. Markopoulou, L. Smolin, "Holography in Quantum Space – Time", hep-th/9910146
L. Smolin, "M Theory as a Matrix Extension of Chern-Simons Theory", Nucl.Phys.B591:227-242,2000, [hep-th/0002009]
Y. Ling, L. Smolin, "Eleven – Dimensional Supergravity as a Constrained Topological Field Theory", Nucl.Phys.B601:191-208,2001, [hep-th/0003285]
L. Smolin, "The Cubic Matrix Model and Duality Between Strings and Loops", [hep-th/0006137]
Y. Ling, L. Smolin, "Holographic Formulation of Quantum Supergravity", Phys.Rev.D63:064010,2001, [hep-th/0009018]
L. Smolin, "The Exceptional Jordan Algebra and the Matrix String", hep-th/0104050

S. Alexander, Y. Ling, L. Smolin, "A Thermal Instability for Positive Brane Cosmological Constant in the Randall-Sundrum Cosmologies", hep-th/0106097

[271] M. Arnsdorf, L. Smolin, "The Maldacena Conjecture and Rehren Duality", hep-th/0106073

[272] J. Maldacena, "The large N Limit of Superconformal Field Theories and Supergravity", Adv.Theor.Math.Phys.2:231-252,1998, Int.J.Theor.Phys.38:1113-1133,1999, [hep-th/9711200]
E. Witten, "Anti-De Sitter Space and Holography", Adv.Theor.Math.Phys.2:253-291,1998, [hep-th/9802150]

[273] K.H. Rehren, "Algebraic Holography", Annales Henri Poincare 1:607-623,2000, [hep-th/9905179]
K.H. Rehren, "Proof of the AdS/CFT Correspondence", in: "Quantum Theory and Symmetries", H.-D. Doebner, et al. (ed.) World Scientific, 2000. pp. 278-284, [hep-th/9910074]
K.-H. Rehren, "Local Quantum Observables in the AdS/CFT Correspondence", Phys.Lett.B493:383-388,2000, [hep-th/0003120]

[274] B. Schroer, "Facts and Fictions about Anti-De Sitter Space-Times with Local Quantum Matter", Commun.Math.Phys.219:57-76,2001, [hep-th/9911100]

[275] B. Hartmann, "Generalized Wick Transform in Dimensionally Reduced Gravity", Diploma Thesis, Humboldt – Universität Berlin, 50 p., November 2001

[276] W. Junker, "Hadamard States, Adiabatic Vacua and the Construction of Physical States for Scalar Quantum Fields on Curved Spacetime", Rev.Math.Phys.8:1091-1159,1996

[277] J. Pullin, R. Gambini, "Making Classical and Quantum Canonical General Relativity Computable Through a Power Series Expansion in the Inverse Cosmological Constant", Phys.Rev.Lett.85:5272-5275,2000, [gr-qc/0008031]
R. Gambini, J. Pullin, "The Large Cosmological Constant Approximation to Classical and Quantum Gravity: Model Examples", Class.Quant.Grav.17:4515-4540,2000, [gr-qc/0008032]

[278] N. Biggs, "Algebraic Graph Theory", Cambridge University Press, 2nd ed., Cambridge, 1993
D. Stauffer, A. Aharony, "Introduction to Percolation Theory", Taylor and Francis, 2nd ed., London, 1994
D. M. Cvetovic, M. Doob, H. Sachs, "Spectra of Graphs", Academic Press, New York, 1979

[279] S. Melosch, H. Nicolai, "New Canonical Variables for $D = 11$ Supergravity", Phys.Lett.B416:91-100,1998, [hep-th/9709227]

[280] C. Beetle, A. Corichi, "Bibliography of Publications Related to Classical and Quantum Gravity in terms of Connection and Loop Variables", gr-qc/9703044